



Atty. Dkt. No. 084335-0182

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Takao ISOGAI et al.
Title: FULL-LENGTH cDNA
Appl. No.: 10/777,288
Filing Date: 02/13/2004
Examiner: Unassigned
Art Unit: 1653

CLAIM FOR CONVENTION PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

The benefit of the filing dates of the following prior foreign applications filed in the following foreign country is hereby requested, and the right of priority provided in 35 U.S.C. § 119 is hereby claimed.

In support of this claim, filed herewith are certified copies of said original foreign applications:


Japanese Patent Application No. 2003-102207 filed February 14, 2003 and
Japanese Patent Application No. 2003-131452 filed May 9, 2003.

A CD-R is attached hereto providing English language translations of domestic priority applications, U.S. SN 60/490,871 and U.S. SN 60/476,172. The CD-R is attached for the benefit of the Examiner and for examination purposes only.

Respectfully submitted,

Date 11-10-2004

FOLEY & LARDNER LLP
Customer Number: 22428
Telephone: (202) 672-5569
Facsimile: (202) 672-5399

 (45,943)
Stephen B. Maebius
Attorney for Applicant
Registration No. 35,264
SEAWA PASSINO FOR

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 5月 9日
Date of Application:

出願番号 特願2003-131452
Application Number:
[ST. 10/C]: [JP 2003-131452]

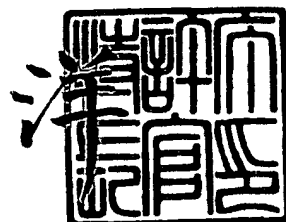
出願人 バイオテクノロジー開発技術研究組合
Applicant(s):

CERTIFIED COPY OF
PRIORITY DOCUMENT

2004年 7月 9日

特許庁長官
Commissioner,
Japan Patent Office

小川



BEST AVAILABLE COPY

出証番号 出証特2004-3059663

- 【書類名】 特許願
- 【整理番号】 BTR-A0303
- 【提出日】 平成15年 5月 9日
- 【あて先】 特許庁長官 殿
- 【国際特許分類】 C12N 15/00
- 【発明者】
- 【住所又は居所】 茨城県稲敷郡阿見町大室 5 1 1 - 1 2
- 【氏名】 磯貝 隆夫
- 【発明者】
- 【住所又は居所】 千葉県木更津市高柳 1 4 8 6 - 1 - E 2 0 5
- 【氏名】 山本 順一
- 【発明者】
- 【住所又は居所】 東京都板橋区氷川町 2 7 - 3 - 4 0 3
- 【氏名】 西川 哲夫
- 【発明者】
- 【住所又は居所】 千葉県木更津市高柳 1 4 8 5 - A 2 0 3
- 【氏名】 五十野 祐子
- 【発明者】
- 【住所又は居所】 東京都墨田区横川 5 - 4 - 3 - 5 1 2
- 【氏名】 杉山 友康
- 【発明者】
- 【住所又は居所】 東京都東久留米市中央町 1 - 2 - 5 - 1 0 2
- 【氏名】 大槻 哲嗣
- 【発明者】
- 【住所又は居所】 千葉県木更津市高柳 1 4 7 3 - 4 - 2 0 2
- 【氏名】 若松 愛
- 【発明者】
- 【住所又は居所】 千葉県木更津市矢那 4 5 0 8 - 1 9 - 2 0 2
- 【氏名】 石井 静子

【発明者】

【住所又は居所】 東京都東大和市桜が丘 3 - 4 4 - 1 4 - 9 - 2 0 4

【氏名】 永井 啓一

【発明者】

【住所又は居所】 埼玉県さいたま市太田窪 1 - 6 - 7

【氏名】 入江 亮太郎

【特許出願人】

【識別番号】 502235522

【氏名又は名称】 バイオテクノロジー開発技術研究組合

【代表者】 理事長 倉内 憲孝

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 全長 cDNA

【特許請求の範囲】

【請求項 1】 下記 (a) から (g) のいずれかに記載のポリヌクレオチド。

(a) 配列番号: 1 ~ 配列番号: 39 のいずれかに記載された塩基配列の蛋白質コード領域を含むポリヌクレオチド。

(b) 配列番号: 40 ~ 配列番号: 78 のいずれかに記載のアミノ酸配列からなる蛋白質をコードする塩基配列を含むポリヌクレオチド。

(c) 配列番号: 40 ~ 配列番号: 78 から選択されたいずれかの配列番号に記載のアミノ酸配列において、1 若しくは複数のアミノ酸が置換、欠失、挿入、および/または付加したアミノ酸配列からなり、前記選択されたアミノ酸配列からなる蛋白質と機能的に同等な蛋白質をコードする塩基配列を含むポリヌクレオチド。

(d) 配列番号: 1 ~ 配列番号: 39 から選択されたいずれかの配列番号に記載の塩基配列からなるポリヌクレオチドとハイブリダイズするポリヌクレオチドであって、前記選択された塩基配列によってコードされる蛋白質と機能的に同等な蛋白質をコードする塩基配列を含むポリヌクレオチド。

(e) (a) から (d) に記載のポリヌクレオチドによってコードされる蛋白質の部分アミノ酸配列をコードする塩基配列を含むポリヌクレオチド。

(f) 配列番号: 1 ~ 配列番号: 39 のいずれかに記載の塩基配列に対して少なくとも 70% の同一性を有する塩基配列を含むポリヌクレオチド。

(g) 配列番号: 1 ~ 配列番号: 39 のいずれかに記載の塩基配列に対して少なくとも 90% の同一性を有する塩基配列を含むポリヌクレオチド。

【請求項 2】 請求項 1 に記載のポリヌクレオチドのいずれか一つによってコードされる蛋白質、またはその部分ペプチド。

【請求項 3】 請求項 2 に記載されたいずれかの蛋白質、またはペプチドに結合する抗体。

【請求項 4】 請求項 2 に記載されたいずれかの蛋白質、またはペプチドと、請求項 3 に記載の抗体とを接触させ、両者の結合を観察する工程を含む、請求項

2 に記載されたいずれかの蛋白質、またはペプチドの免疫学的測定方法。

【請求項 5】請求項 1 に記載されたポリヌクレオチドのいずれか一つを含むベクター。

【請求項 6】請求項 1 に記載のポリヌクレオチド、または請求項 5 に記載のベクターを保持する形質転換体。

【請求項 7】請求項 1 に記載されたポリヌクレオチドのいずれか一つ、または請求項 5 に記載のベクターを発現可能に保持する形質転換体。

【請求項 8】請求項 7 に記載の形質転換体を培養し、発現産物を回収する工程を含む、請求項 2 に記載されたいずれかの蛋白質またはペプチドの製造方法。

【請求項 9】配列番号：1～配列番号：39 のいずれかに記載された塩基配列、またはその相補鎖に相補的な塩基配列からなる 15 ヌクレオチド以上の鎖長を持つオリゴヌクレオチド。

【請求項 10】請求項 9 に記載のオリゴヌクレオチドからなる、ポリヌクレオチド合成用プライマー。

【請求項 11】請求項 9 に記載のオリゴヌクレオチドからなる、ポリヌクレオチドの検出用プローブ。

【請求項 12】下記 (a) から (c) のいずれかに記載のポリヌクレオチド。

(a) 請求項 1 に記載のポリヌクレオチドの転写産物と相補的な塩基配列を有するアンチセンスポリヌクレオチド

(b) 請求項 1 に記載のポリヌクレオチドの転写産物を特異的に開裂するリボザイム活性を有するポリヌクレオチド

(c) 宿主細胞における発現時に、RNAi 効果により、請求項 1 に記載のポリヌクレオチドの発現を抑制するポリヌクレオチド

【請求項 13】次の工程を含む、請求項 1 に記載のポリヌクレオチドの検出方法。

a) 標的ポリヌクレオチドと請求項 9 に記載のオリゴヌクレオチドを、ハイブリダイゼーションが可能な条件下でインキュベートする工程、

b) 標的ポリヌクレオチドと請求項 9 に記載のオリゴヌクレオチドのハイブリダ

イゼーションを検出する工程。

【請求項 1 4】 配列番号： 1 ～ 配列番号： 3 9 のいずれかに記載された塩基配列および／または配列番号： 4 0 ～ 配列番号： 7 8 のいずれかに記載のアミノ酸配列から選択された少なくとも 1 つの配列情報を含むポリヌクレオチドおよび／または蛋白質データベース。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、新規な蛋白質をコードするポリヌクレオチド、このポリヌクレオチドによってコードされる蛋白質、及びそれらの新規な用途に関する。

【 0 0 0 2 】

【従来の技術】

現在、世界的なレベルで様々な生物のゲノム配列の解明とその解析が進められている。既に40種類を越える原核微生物、下等真核生物の出芽酵母、多細胞性真核生物である線虫、高等植物であるシロイヌナズナ等で、その全ゲノム配列が決定された。30億塩基対といわれるヒトのゲノムについては、世界的な協力体制のもとでその解析が進められて2001年にドラフト配列が公開された。そして、2 0 0 3 年にはその全構造が明らかになり、公開されようとしている。ゲノム配列を明らかにする目的は、全ての遺伝子の機能や制御、あるいは遺伝子間、蛋白質間、細胞間さらには個体間における相互作用のネットワークとして複雑な生命現象を理解するところにある。種々の生物種のゲノム情報から生命現象を解明していくことは、単に学術分野における研究課題として重要であるのみならず、そこで得られる研究成果をいかに産業上の応用へと発展させていくかという点で、その社会的な意義も大きい。

【 0 0 0 3 】

ところが単にゲノム配列を決定しただけでは、全ての遺伝子の機能を明らかにできるわけではない。例えば酵母では、ゲノム配列から推定された約6,000の遺伝子の約半数しか、その機能を推定できなかった。一方、ヒトには約3～4万種類の遺伝子が存在すると推測されており、さらにオルタナティブスプライシングに

よるバリエーションも考慮に入れると10万種以上のmRNAが存在すると言われている。そこで、ゲノム配列から明らかにされてくる膨大な量の新しい遺伝子の機能を、迅速かつ効率的に解明していくための「ハイスループット遺伝子機能解析システム」の確立が、強く望まれている。

【0 0 0 4】

真核生物のゲノム配列では、多くの場合、一つの遺伝子がイントロンによって複数のエクソンに分断されている。そのため、ゲノム配列情報だけからそこにコードされる蛋白質の構造を正確に予測するには、多くの問題がある。一方、イントロンが除かれたmRNAから作製されるcDNAでは、蛋白質のアミノ酸配列の情報が一つの連続した配列情報として得られるため、容易にその一次構造を明らかにすることが可能である。ヒトのcDNAの研究では、これまでに300万を越えるEST (Expression Sequence Tags) データがパブリックドメインに公開されており、それらはヒトの全遺伝子の80%以上をカバーしているものと推定されている。

これらの情報は、ヒト遺伝子構造の解明やゲノム配列におけるエクソン領域の予測、あるいはその発現プロファイルの推定など、様々な角度から利用されている。ところが、これらのヒトEST情報の多くはcDNAの3' 末端側近傍に集中しているため、特にmRNAの5' 末端近傍の情報が極端に不足している状況にある。また、ヒトcDNAの中で全長でコードされている蛋白質の配列が予測されているmRNAは約1万5千種類程度である。

【0 0 0 5】

完全長cDNAでは、その5' 末端配列からゲノム配列上でのmRNA転写開始点が特定できる上、その配列の中に含まれるmRNAの安定性や翻訳段階での発現制御に関わる因子の解析が可能である。また、翻訳開始点であるATGコドンをも5' 側に含むことから、正しいフレームで蛋白質への翻訳を行うことができる。したがって、適当な遺伝子発現系を適用することで、そのcDNAがコードする蛋白質を大量に生産したり、蛋白質を発現させてその生物学的活性を解析することも可能になる。このように、完全長cDNAの解析からはゲノム配列解析を相補する重要な情報が得られる。また、発現可能な全長cDNAクローンは、その遺伝子の機能の実証的な解析や産業分野での応用への展開において、その重要性はきわめて高い。

【0 0 0 6】

したがって新規なヒト全長cDNAが単離されれば、それらの遺伝子が関与している種々の疾患に対する医薬品開発に利用され得る。これらの遺伝子がコードする蛋白質はそれ自身に医薬品としての有用性を期待できる。したがって、新規なヒト蛋白質をコードするcDNAの全長を取得することには大きな意義がある。

特にヒト分泌蛋白質、または膜蛋白質には、そのものがティッシュプラスミノーゲンアクチベーター (TPA)のように、医薬品として有用なものや、膜レセプターのように医薬品の標的蛋白質になりうるものが多い。また、シグナル伝達関連蛋白質 (Protein kinase等)、糖蛋白質関連蛋白質、転写関連蛋白質等は疾患との関係が解明されている遺伝子が多い。更に疾患関連蛋白質の遺伝子は、ヒト遺伝子と疾患の関係が解明されている遺伝子が多い遺伝子群である。

したがって、ヒトにおいて分離が進んでいない新規な全長cDNAを提供する意義は大きい。中でも、分泌・膜蛋白質をコードする蛋白質をコードするcDNAは、蛋白質自身に医薬品としての有用性を期待できること、あるいは疾患に関連する遺伝子を多く含む可能性のあることから、未知のcDNAの分離が望まれている。あるいは、疾患との関連性が強いと予測される遺伝子群である、シグナル伝達蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質、そして疾患関連蛋白質をコードする遺伝子も、治療のための標的分子として、またこれらの蛋白質自身に医薬品としての有用性を期待できる。したがって、これらの蛋白質をコードするcDNAの全長を明らかにすることには大きな意義がある。

【0 0 0 7】

【発明が解決しようとする課題】

本発明は、新規な蛋白質をコードするポリヌクレオチド、該ポリヌクレオチドによってコードされる蛋白質、並びにその用途の提供を課題としている。

【0 0 0 8】

【課題を解決するための手段】

本発明者らは、オリゴキャップ法[K. Maruyama and S. Sugano, Gene, 138: 171-174 (1994); Y. Suzuki et al., Gene, 200: 149-156 (1997)]を改良した方法(WO 01/04286)で作成した全長率の非常に高いヒトcDNAライブラリーから、全

長cDNAクローンであると予測される、ヒト全長cDNAを効率よくクローニングする方法を開発した。次いで、この方法で取得した全長率の高いcDNAクローンの塩基配列を主に5'側から、また必要に応じ3'側からも決定した。

更に、得られたクローンのうち新規で全長と予測される代表的クローンについて全長塩基配列を決定した。得られた全長塩基配列について、以下に示すデータベースに対してBLASTを用いた相同性検索を行った。本発明の相同性検索は、cDNAの全コード領域を含む全長cDNAの情報に基づいて行われているので、蛋白質のあらゆる部分に対する相同性を解析することができる。したがって本発明においては、相同性検索の信頼性が飛躍的に向上している。

[1] SwissProt (http://www.ebi.ac.uk/ebi_docs/SwissProt_db/swisshome.html)、

[2] GenBank(<http://www.ncbi.nlm.nih.gov/web/GenBank/>)、

[3] nr(重複を除いたGenBankの塩基配列のコーディングシーケンス (CDS) データ、

SwissProt、

PDB(<http://www.rcsb.org/pdb/index.html>)、

PIR(<http://pir.georgetown.edu/pirwww/pirhome.shtml>)、

PRF(<http://www.prf.or.jp/en/>)、を組み合わせで構築されている蛋白質データベース)、および

[4] RefSeq(<http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html>)

【 0 0 0 9 】

また得られた5'側の塩基配列をもとに構築した大規模cDNAデータベースを解析して、全長塩基配列を決定したクローンの遺伝子発現プロファイルを解析した。本発明者らは、これらの解析の結果に基づいて、本発明の遺伝子の有用性を明らかにした。

本発明においては、全長塩基配列情報に基づくin silicoでの発現プロファイルの解析によって遺伝子の機能を明らかにしている。発現頻度解析に用いた発現プロファイルは、十分量の断片配列のデータベースに基づいて解析した。この発現プロファイルに、本発明で得られた多くのcDNAクローンの全長塩基配列を照合

することによって、発現頻度解析を行った。このように、量的に十分な解析母体（発現プロファイル）に対して、幅広い遺伝子の全長塩基配列を照合したことによって、信頼性の高い解析が可能となった。すなわち、本発明の全長配列を利用した発現頻度解析の結果は、あるcDNAライブラリーの由来となった組織や細胞の遺伝子発現頻度をより正確に反映していると言えることができる。つまり、本発明のcDNAの全長塩基配列情報によって、信頼性の高い発現頻度解析が可能となったことを意味している。

【0010】

本発明における全長cDNAクローンは、[1]オリゴキャップ法による全長率の高いcDNAライブラリーの作成、および[2] 5'末端側の配列をアセンブルし、その結果形成されたクラスターの中でも最も全長らしいもの（5'側に長いものが多い）を選択する、という方法を組み合わせて取得した、全長cDNAクローンである。しかし、本発明が提供するポリヌクレオチドの5'末端および3'末端の配列をもとに設計されたプライマーを利用すれば、必ずしもこのような特殊な方法を用いずとも、容易に全長cDNAの取得が可能となる。なお、発現可能なcDNAを取得するためのプライマー設計は、ポリヌクレオチドの5'末端および3'末端配列を用いることに限定されない。

【0011】

すなわち本発明は、次のポリヌクレオチドと、このポリヌクレオチドによってコードされる蛋白質に関する。

[1] 下記（a）から（g）のいずれかに記載のポリヌクレオチド。

（a）配列番号：1～配列番号：39のいずれかに記載された塩基配列の蛋白質コード領域を含むポリヌクレオチド。

（b）配列番号：40～配列番号：78のいずれかに記載のアミノ酸配列からなる蛋白質をコードする塩基配列を含むポリヌクレオチド。

（c）配列番号：40～配列番号：78から選択されたいずれかの配列番号に記載のアミノ酸配列において、1若しくは複数のアミノ酸が置換、欠失、挿入、および／または付加したアミノ酸配列からなり、前記選択されたアミノ酸配列からなる蛋白質と機能的に同等な蛋白質をコードする塩基配列を含むポリヌクレオ

チド。

(d) 配列番号: 1 ~ 配列番号: 39 から選択されたいずれかの配列番号に記載の塩基配列からなるポリヌクレオチドとハイブリダイズするポリヌクレオチドであって、前記選択された塩基配列によってコードされる蛋白質と機能的に同等な蛋白質をコードする塩基配列を含むポリヌクレオチド。

(e) (a) から (d) に記載のポリヌクレオチドによってコードされる蛋白質の部分アミノ酸配列をコードする塩基配列を含むポリヌクレオチド。

(f) 配列番号: 1 ~ 配列番号: 39 のいずれかに記載の塩基配列に対して少なくとも 70% の同一性を有する塩基配列を含むポリヌクレオチド。

(g) 配列番号: 1 ~ 配列番号: 39 のいずれかに記載の塩基配列に対して少なくとも 90% の同一性を有する塩基配列を含むポリヌクレオチド。

[2] [1] に記載のポリヌクレオチドのいずれか一つによってコードされる蛋白質、またはその部分ペプチド。

[3] [2] に記載されたいずれかの蛋白質、またはペプチドに結合する抗体。

[4] [2] に記載されたいずれかの蛋白質、またはペプチドと、[3] に記載の抗体とを接触させ、両者の結合を観察する工程を含む、[2] に記載されたいずれかの蛋白質、またはペプチドの免疫学的測定方法。

[5] [1] に記載されたポリヌクレオチドのいずれか一つを含むベクター。

[6] [1] に記載のポリヌクレオチド、または [5] に記載のベクターを保持する形質転換体。

[7] [1] に記載されたポリヌクレオチドのいずれか一つ、または [5] に記載のベクターを発現可能に保持する形質転換体。

[8] [7] に記載の形質転換体を培養し、発現産物を回収する工程を含む、[2] に記載されたいずれかの蛋白質またはペプチドの製造方法。

[9] 配列番号: 1 ~ 配列番号: 39 のいずれかに記載された塩基配列、またはその相補鎖に相補的な塩基配列からなる 15 ヌクレオチド以上の鎖長を持つオリゴヌクレオチド。

[10] [9] に記載のオリゴヌクレオチドからなる、ポリヌクレオチド合成用プライマー。

〔11〕〔9〕に記載のオリゴヌクレオチドからなる、ポリヌクレオチドの検出用プローブ。

〔12〕下記（a）から（c）のいずれかに記載のポリヌクレオチド。

（a）〔1〕に記載のポリヌクレオチドの転写産物と相補的な塩基配列を有するアンチセンスポリヌクレオチド

（b）〔1〕に記載のポリヌクレオチドの転写産物を特異的に開裂するリボザイム活性を有するポリヌクレオチド

（c）宿主細胞における発現時に、RNAi効果により、〔1〕に記載のポリヌクレオチドの発現を抑制するポリヌクレオチド

〔13〕次の工程を含む、〔1〕に記載のポリヌクレオチドの検出方法。

a) 標的ポリヌクレオチドと〔9〕に記載のオリゴヌクレオチドを、ハイブリダイゼーションが可能な条件下でインキュベートする工程、

b) 標的ポリヌクレオチドと〔9〕に記載のオリゴヌクレオチドのハイブリダイゼーションを検出する工程。

〔14〕配列番号：1～配列番号：39のいずれかに記載された塩基配列および／または配列番号：40～配列番号：78のいずれかに記載のアミノ酸配列から選択された少なくとも1つの配列情報を含むポリヌクレオチドおよび／または蛋白質データベース。

【0012】

本発明において、ポリヌクレオチドとはDNAやRNAのようにヌクレオチドが多数重合した分子を意味する。重合するヌクレオチドの数は特に制限されない。ポリヌクレオチドの重合度が比較的低い場合には特にオリゴヌクレオチドとも表現するが、これも本発明のポリヌクレオチドに含まれる。本発明のポリヌクレオチド、またはオリゴヌクレオチドは、天然のものであることもできるし、化学的に合成されたものであることもできる。あるいはまた、鋳型となるDNAをもとにPCRのような酵素的な反応によって合成されたものであっても良い。更に本発明のポリヌクレオチドは、化学的に修飾されたものであっても良い。また本発明には、1本鎖ポリヌクレオチドのみならず、2本鎖ポリヌクレオチドも含まれる。本明細書、特に請求項において、単にポリヌクレオチドと記載するときには、1本鎖ポ

リヌクレオチドのみならず2本鎖ポリヌクレオチドをも指すものとする。2本鎖ポリヌクレオチドを意味するときには、一方の鎖のみについての塩基配列が記載されることになるが、センス鎖の塩基配列に基づいてその相補鎖の塩基配列は必然的に規定される。

【0 0 1 3】

本発明によって提供されるcDNAはいずれも全長cDNAである。本発明における全長cDNAとは、そのcDNAの翻訳開始点となるATGコドンを含むことを意味する。天然のmRNAが蛋白質コード領域の上流や下流に本来備えている非翻訳領域の有無は問わない。また本発明の全長cDNAは、望ましくは終止コドンを含む。

【0 0 1 4】

【発明の実施の形態】

本発明のすべてのクローン（39クローン）は新規で、蛋白質の全長をコードするものである。またすべてのクローンは、全長性の高いオリゴキャップ法で取得されたcDNAであり、その5'末端配列をGenBank、UniGene(Human) (<http://www.ncbi.nlm.nih.gov/UniGene>) データベースの"complete cds"表記のあるmRNA配列に対して、BLAST [S. F. Altschul, W. Gish, W. Miller, E. W. Myers & D. J. Lipman, J. Mol. Biol., 215: 403-410 (1990); W. Gish & D. J. States, Nature Genet., 3: 266-272 (1993)]により相同性検索を行って選別された、ヒトmRNAに対して同一でない（すなわち新規である）クローンであり、またアセンブルを行った結果形成されたクラスターのメンバーのうちでより全長性が高いと思われるクローンである。クラスター中で全長性が高いと思われる塩基配列は、5'側に長いものが最も多かった。

【0 0 1 5】

本発明の全ての全長cDNAは、5'末端配列と3'末端配列に基づいて設計されたプライマーセット、あるいは5'末端配列に基づいて設計されたプライマーとポリA配列に対応するオリゴdT配列とで構成されるプライマーセットを使用し、PCR (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publishing. John Wiley & Sons Section 6.1-6.4) 等の手法を用いることにより合成することができる。表1に、本発明の全長cDNA 39クローンのクローン名とその全長

塩基配列を示す配列番号、および全長塩基配列から推定されるCDS部位と翻訳されたアミノ酸を示す配列番号を示す。CDS部位の存在位置については、「DDBJ/EMBL/GenBank Feature Table Definition」 (<http://www.ncbi.nlm.nih.gov/collab/FT/index.html>) の規則に従って記載した。開始位置番号はメチオニンをコードする塩基である「ATG」の1文字目であり、終了位置番号はストップコドンの3文字目である。これを「..」で挟んで記載した。ただし、終止コドンが現れないクローンについては上記規則に則って「>」を用いて終了位置を記載した。

【0 0 1 6】

【表 1】

クローン名	塩基配列 番号	CDS位置	アミノ酸 配列番号
ADRL2011190	1	56..2113	40
BRACE3002184	2	199..1890	41
BRACE3026993	3	143..4903	42
BRACE3046450	4	361..4356	43
BRAMY3008096	5	602..1396	44
BRAMY3016953	6	4547..5188	45
BRAWH3013197	7	233..2623	46
BRAWH3028645	8	256..1938	47
BRAWH3046240	9	2257..3441	48
BRCAN2019772	10	5..1273	49
BRHIP3030064	11	3039..3347	50
BRHIP3038037	12	122.. 739	51
BRTHA3004432	13	1988..2629	52
BRTHA3024233	14	1440..2039	53
CTONG2002832	15	346..2148	54
CTONG2003764	16	1438..2169	55

PLACE7013963	17	2326..>5202	56
PROST2010326	18	298..1869	57
TBAES2004105	19	135..>1864	58
TBAES2007379	20	70..2118	59
TBAES2007481	21	53..1162	60
TBAES2008133	22	389..2719	61
TESTI2043585	23	310..1296	62
TESTI2046536	24	111.. 533	63
TESTI4002988	25	597..2687	64
TESTI4005158	26	952..2961	65
TESTI4005500	27	249..2336	66
TESTI4052089	28	1517..2038	67
THYMU3024602	29	31..2634	68
THYMU3044175	30	1101..2054	69
THYMU3046350	31	282..1787	70
TLIVE2007192	32	350.. 730	71
TRACH1000193	33	998..2956	72
TRACH3019290	34	152..1738	73
TRACH3021066	35	2485..3792	74
UTERU2017492	36	80.. 808	75
UTERU2025415	37	47.. 379	76
UTERU2036507	38	276..1691	77
UTERU3020583	39	126.. 575	78

【 0 0 1 7 】

すなわち前記表 1 中の配列番号に示した本発明のポリヌクレオチドの塩基配列に基づいて、当該ポリヌクレオチドを合成するためのプライマーをデザインすることができる。なお全長cDNAの合成を目的とするとき、3'側のプライマーとしてはオリゴdTプライマーを用いることもできる。プライマーの長さは、通常、15bp

～100bp、好ましくは15bp～35bpの鎖長を有する。後に述べるLA PCRに用いる場合には、25～35bpの長さとするると良い結果を得ることができる。

【0018】

目的とする塩基配列に基づいて、特異的な増幅を可能とするプライマーを設計する手法は公知である（Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons Section 6.1-6.4）。5'側の塩基配列に基づいてプライマーを設計する際には、原則として増幅生成物が翻訳開始点を含むようにする。したがって、たとえば5'側のプライマーを5'側の非翻訳領域(5' UTR)の塩基配列に基づいて設定する場合には、対象となるcDNAに対する特異性を保証できる限り、任意の部分を5'側のプライマーとして選択することができる。

【0019】

全長cDNAを合成する場合には、その増幅対象塩基配列は長いものでは数千bpにも及ぶ。しかしLA PCR(Long and Accurate PCR)等を利用することにより、このような長い領域の合成は可能である。長いDNAの合成には、LA PCRを利用するのが有利である。LA PCRでは、3'→5'エキソヌクレアーゼ活性を持った特殊なDNAポリメラーゼを用いることにより、誤って取りこまれた塩基を除去できる。そのため、長い塩基配列であっても正確に相補鎖合成を進めることができるのである。LA PCRを利用すれば、望ましい条件においては、20 kb以上の増幅が可能とされている。（林健志、実験医学別冊・PCRの最新技術、羊土社, 1996）

【0020】

本発明の全長cDNAを合成するための鋳型となるDNAには、様々な方法によって調製したcDNAライブラリーを利用することができる。本発明における全長cDNAクローンは、[1]オリゴキャップ法による全長率の非常に高いcDNAライブラリーの作製、および[2] 5'末端側の配列をアセンブルし、その結果形成されたクラスターの中でも最も全長らしいクローン（5'側に長いものが多い）を選択する、という方法を組み合わせて取得した、より全長である確率の高いクローンである。

しかし、本発明によって提供される全長塩基配列に基づいてデザインされるプライマーを利用すれば、必ずしもこのような特殊な方法を用いずとも、容易に全

長cDNAの取得が可能となる。

【0021】

すなわち、公知の方法によって調製されたcDNAライブラリー、あるいは市販のcDNAライブラリーは、全長mRNAをまったく含まないものではなく、その割合が非常に低い。したがって、通常のクローニング方法では、これらのライブラリーから全長cDNAクローンを直接スクリーニングすることは困難である。しかし、本発明によって新規な全長cDNAの塩基配列が明らかにされた。全長塩基配列が与えられれば、PCRのような酵素的な合成方法を利用して目的とする全長cDNAを合成することが可能である。ただし、より確実に全長cDNAの合成を行うには、たとえばオリゴキャップ法等によって合成された全長率の高いcDNAライブラリーの使用が望ましいことは言うまでもない。

【0022】

本発明の全長cDNAクローンの5'-端を含む塩基配列を利用すれば、ゲノム上のプロモーターを含む転写制御領域を単離することが可能となる。既にヒトゲノムの90%以上をカバーするラフドラフト（精度が少し低いヒトゲノム配列解析）が報告されている（Nature, Vol.409, 814-823, 2001）。さらに、2003年にはヒト全ゲノム配列解析が完了する計画になっている。長いイントロンの存在するヒトゲノムより転写開始点を解析ソフトで解析することは大きな困難がともなう。しかし、本発明の全長cDNAクローンの5'-端を含む塩基配列を用いれば、全長cDNAの5'-端を含む塩基配列からゲノム配列上でのmRNA転写開始点を容易に特定できるため、転写開始点上流配列の中に含まれるプロモーターを含む転写制御に関わるゲノム領域を取得することが容易となる。

【0023】

本発明の全長cDNAによってコードされる蛋白質は、組み換え蛋白質として、また天然の蛋白質として調製することが可能である。組み換え蛋白質は、例えば、後述するように本発明の蛋白質をコードするDNAを挿入したベクターを適当な宿主細胞に導入し、形質転換体内で発現した蛋白質を精製することにより調製することが可能である。一方、天然の蛋白質は、例えば、後述する本発明の蛋白質に対する抗体を結合したアフィニティーカラムを利用して調製することができる（

Current Protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. Jhon Wiley & Sons Section 16.1-16.19)。アフィニティー精製に用いる抗体は、ポリクローナル抗体であってもモノクローナル抗体であってもよい。また、インビトロトランスレーション（例えば、「On the fidelity of mRNA translation in the nuclease-treated rabbit reticulocyte lysate system. Dasso, M. C., Jackson, R. J. (1989) Nucleic Acids Res. 17:3129-3144」参照）などにより本発明の蛋白質を調製することも可能である。

【0024】

前記のようにして明らかにされた本発明による蛋白質の活性に基づいて、本発明の蛋白質と機能的に同等な蛋白質を得ることができる。ある蛋白質が本発明の蛋白質と機能的に同等であるかどうかは、本発明の蛋白質が備える生物学的な活性を指標として、該活性をある蛋白質が有するかどうかを調べることによって確認することができる。

【0025】

これら本発明において同定された蛋白質と機能的に同等な蛋白質は、当業者であれば、例えば、蛋白質中のアミノ酸配列に変異を導入する方法（例えば、部位特異的変異誘発法(Current Protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. Jhon Wiley & Sons Section 8.1-8.5))を利用して調製することができる。また、このような蛋白質は、自然界におけるアミノ酸の変異により生じることもある。本発明には、このように本実施例において同定された蛋白質と同等の機能を有する限り、そのアミノ酸配列（表1）において1もしくは複数のアミノ酸が置換、欠失、挿入および／または付加された蛋白質も含まれる。

【0026】

蛋白質におけるアミノ酸の変異数や変異部位は、その機能が保持される限り制限はない。変異数は、典型的には、30%以内、または20%以内、または10%以内であり、好ましくは全アミノ酸の5%以内、または3%以内であり、さらに好ましくは全アミノ酸の2%以内であり、更に好ましくは全アミノ酸の1%以内である。あるいは本発明には複数のアミノ酸として数個のアミノ酸の変異を置

換する場合が含まれる。数個とは、たとえば5、更には4または3、あるいは2、更には1のアミノ酸を言う。

置換されるアミノ酸は、蛋白質の機能の保持の観点から、置換前のアミノ酸と似た性質を有するアミノ酸であることが好ましい。例えば、Ala、Val、Leu、Ile、Pro、Met、Phe、Trpは、共に非極性アミノ酸に分類されるため、互いに似た性質を有すると考えられる。また、非荷電性としては、Gly、Ser、Thr、Cys、Tyr、Asn、Glnが挙げられる。また、酸性アミノ酸としては、AspおよびGluが挙げられる。また、塩基性アミノ酸としては、Lys、Arg、Hisが挙げられる。

【0027】

また、本実施例において同定された蛋白質と機能的に同等な蛋白質は、当業者に周知のハイブリダイゼーション技術あるいは遺伝子増幅技術を利用して単離することも可能である。即ち、当業者であれば、ハイブリダイゼーション技術 (Current Protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. Jhon Wiley & Sons Section 6.3-6.4) を用いて本実施例において同定された蛋白質をコードするDNAの塩基配列 (表1) またはその一部をもとにこれと相同性の高いDNAを単離して、該DNAから機能的に同等な蛋白質を得ることは、通常行いうることである。本発明には、本実施例において同定された蛋白質と同等の機能を有する限り、これら蛋白質をコードするDNAとハイブリダイズするDNAによりコードされる蛋白質も含まれる。機能的に同等な蛋白質を単離する生物としては、例えば、ヒト、マウス、ラット、ウサギ、ブタ、ウシ等の脊椎動物が挙げられるが、これらに制限されない。

【0028】

機能的に同等な蛋白質をコードするDNAを単離するためのハイブリダイゼーションの条件は、洗浄条件として通常「1xSSC、0.1% SDS、37℃」程度であり、より厳しい条件としては「0.5xSSC、0.1% SDS、42℃」程度であり、さらに厳しい条件としては「0.1xSSC、0.1% SDS、65℃」程度を示すことができる。あるいは、次のような条件を本発明におけるハイブリダイゼーションの条件として示すこともできる。すなわち、ハイブリダイゼーションを「6 x SSC、40 % ホルムアミド、25 °C」、洗浄を「1 x SSC、55 °C」で行う条件を用いることができる。よ

り好ましい条件としては、ハイブリダイゼーションを「6 x SSC、40 % ホルムアミド、37 °C」、洗浄を「0.2 x SSC、55 °C」で行う条件が挙げられる。さらに好ましい条件としては、ハイブリダイゼーションを「6 x SSC、50 % ホルムアミド、37 °C」、洗浄を「0.1 x SSC、62 °C」で行う条件を用いることができる。ハイブリダイゼーションの条件が厳しくなるほどプローブ配列と高い相同性を有するDNAの単離を期待しうる。したがって、ハイブリダイゼーションはストリンジェントな条件下で行うことが望ましい。本明細書においてストリンジェントな条件としては、洗浄のための条件としてたとえば「0.5xSSC、0.1% SDS、42°C」程度を示すことができる。あるいは、ハイブリダイゼーションを「6 x SSC、40 % ホルムアミド、37 °C」、洗浄を「0.2 x SSC、55 °C」で行う条件をストリンジェントな条件として示すこともできる。

【0 0 2 9】

なお、当業者であれば、SSCの希釈率、ホルムアミド濃度、温度などの諸条件を適宜選択することで、上記の条件と同様のストリンジェンシーのハイブリダイゼーション条件を実現することができる。

ただし、上記SSC、SDSおよび温度の条件の組み合わせは例示であり、当業者であれば、ハイブリダイゼーションのストリンジェンシーを決定する上記若しくは他の要素（例えば、プローブ濃度、プローブの長さ、ハイブリダイゼーション反応時間など）を適宜組み合わせることにより、上記と同様のストリンジェンシーを実現することが可能である。

【0 0 3 0】

このようなハイブリダイゼーション技術を利用して単離される蛋白質は表1に記載の本発明の蛋白質と比較して、通常、そのアミノ酸配列において高い相同性を有する。本発明は、請求項1（a）に記載の塩基配列に対して高い同一性を有する塩基配列を含むポリヌクレオチドを包含する。また本発明は、請求項1（b）に記載したポリヌクレオチドがコードするアミノ酸配列に対して高い同一性を有するアミノ酸配列を含む蛋白質、またはペプチドを包含する。高い同一性とは、少なくとも40%以上、好ましくは60%以上、さらに好ましくは70%以上の配列の同一性を指す。あるいはより望ましくは、90%以上、または93%以

上、あるいは95%以上、更には97%以上、そして99%以上の同一性を言う。同一性は、BLAST検索アルゴリズムを用いて決定することができる。

【0031】

本発明におけるアミノ酸配列や塩基配列の相同性は、Karlin and Altschul によるアルゴリズムBLAST (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993)によって決定することができる。このアルゴリズムに基づいて、blastnやblastxと呼ばれるプログラムが開発されている(Altschul et al. J. Mol. Biol.215:403-410, 1990)。BLASTに基づいてblastnによって塩基配列を解析する場合には、パラメーターはたとえばscore = 100、wordlength = 12とする。また、BLASTに基づいてblastxによってアミノ酸配列を解析する場合には、パラメーターはたとえば score = 50、wordlength = 3とする。BLASTとGapped BLASTプログラムを用いる場合には、各プログラムのデフォルトパラメーターを用いる。これらの解析方法の具体的な手法は公知である(<http://www.ncbi.nlm.nih.gov>)。

【0032】

また、遺伝子増幅技術(PCR) (Current protocols in Molecular Biology ed it. Ausubel et al. (1987) Publish. John Wiley & Sons Section 6.1-6.4)を用いて、本実施例において同定された塩基配列(表1)の一部をもとにプライマーを設計し、これらDNA配列またはその一部と相同性の高いDNA断片を単離して、これをもとに本実施例において同定された蛋白質と機能的に同等な蛋白質を得ることも可能である。

【0033】

本発明はまた、表1に示した配列番号に記載の塩基配列からなるポリヌクレオチド、またはその相補鎖に相補的な塩基配列からなる少なくとも15ヌクレオチドを含むポリヌクレオチドを提供する。ここで「相補鎖」とは、A:T、G:Cの塩基対からなる2本鎖DNAの一方の鎖に対する他方の鎖を指す。また、「相補的」とは、少なくとも15個の連続したヌクレオチド領域で完全に相補配列である場合に限られず、少なくとも70%、好ましくは少なくとも80%、より好ましくは90%、さらに好ましくは95%以上の塩基配列上の相同性を有すればよい。塩基配列の相同性は、本明細書に記載したアルゴリズムにより決定することができる。

このようなポリヌクレオチドは、本発明の蛋白質をコードするポリヌクレオチドを検出、単離するためのプローブとして、また、本発明のDNAを増幅するためのプライマーとして利用することが可能である。プライマーとして用いる場合には、通常、15bp～100bp、好ましくは15bp～35bpの鎖長を有する。また、プローブとして用いる場合には、本発明のポリヌクレオチドの少なくとも一部若しくは全部の配列を有し、少なくとも15bpの鎖長のDNAが用いられる。プライマーとして用いる場合、3'側の領域は相補的である必要があるが、5'側には制限酵素認識配列やタグなどを付加することができる。

【 0 0 3 4 】

また、本発明のポリヌクレオチドには、表1に示した配列番号に記載されたアミノ酸配列からなる本発明の蛋白質の発現を抑制するためのアンチセンスポリヌクレオチドが含まれる。アンチセンスポリヌクレオチドは、アンチセンス効果を引き起こすために、少なくとも15bp以上、たとえば50bp以上、好ましくは100bp以上、さらに好ましくは500bp以上の鎖長を有し、通常、3000bp以内、好ましくは2000bp以内の鎖長を有する。このようなアンチセンスDNAには、本発明の蛋白質の異常（機能異常や発現異常）などに起因した疾患の遺伝子治療への応用も考えられる。該アンチセンスDNAは、例えば、本発明の蛋白質をコードするDNA（例えば、配列番号：1～配列番号：39に記載の塩基配列）の配列情報を基にホスホロチオエート法（Stein, 1988 Physicochemical properties of phosphorothioate oligodeoxynucleotides. Nucleic Acids Res 16, 3209-21 (1988)）などにより調製することが可能である。

【 0 0 3 5 】

その他、本発明のポリヌクレオチドの塩基配列に基づいてデザインすることができる、リボザイム、あるいはRNA interference (RNAi)効果により本発明のポリヌクレオチドの発現を抑制することができるポリヌクレオチドも本発明に含まれる。

リボザイムは、本発明のポリヌクレオチドのアンチセンス配列と、触媒作用に必要な触媒ユニットを構成する塩基配列とで構成されるポリヌクレオチドである。リボザイムを構成するアンチセンス配列は、当該リボザイムの触媒ユニットの

構造に合わせて、適宜選択することができる。一方リボザイムの触媒ユニットは、公知である。たとえば、ハンマーヘッド型リボザイム(Rossi et al. (1991) *Pharmac. Ther.* 50: 245-254)やヘアピン型のリボザイム(Hampel et al. (1990) *Nucl. Acids Res.* 18: 299-304, and U.S. Pat. No. 5,254,678)が、塩基配列特異的な切断作用を有することが知られている。これらのリボザイムは、アンチセンス配列がハイブリダイズするポリヌクレオチドの特定の位置を、その触媒作用によって切断することができる。

【0036】

例えば、ハンマーヘッド型リボザイムの自己切断ドメインは、G13U14C15という配列のC15の3'側を切断する。ハンマーヘッド型リボザイムの活性にはU14とA9との塩基対形成が重要とされ、C15の代わりにA15またはU15でも切断できることが示されている(Koizumi M, et al.: *FEBS Lett* 228: 228, 1988)。基質結合部位が標的部位近傍のRNA配列と相補的なリボザイムを設計すれば、標的RNA中のUC、UUまたはUAという配列を認識する制限酵素的なRNA切断リボザイムを作出することができる(Koizumi M, et al.: *FEBS Lett* 239: 285, 1988、小泉誠および大塚栄子: *タンパク質核酸酵素* 35: 2191, 1990、Koizumi M, et al.: *Nucl Acids Res* 17: 7059, 1989)。例えば、本発明のDNA(配列番号: 1から配列番号: 39)中には、標的となり得る部位が複数存在する。アンチセンス配列を選択することにより、本発明のポリヌクレオチドの任意の位置をリボザイムで切断することができる。

【0037】

リボザイムは、好ましくはRNAによって構成される。このようなリボザイムは、化学的に合成することもできるし、あるいは酵素的な反応によって製造することもできる。RNAの化学的な合成方法は公知である。あるいは、リボザイムをコードするDNAをRNAポリメラーゼによって転写することによって、リボザイムを生成することもできる。転写によってリボザイムを生成するときには、リボザイムをコードするDNAは、RNAポリメラーゼが認識するプロモーターの下流に配置される。RNAポリメラーゼとしては、T7 RNAポリメラーゼやSP6RNAポリメラーゼを用いることができる。更に、リボザイムをコードするDNAを適当なベクターに組み

込んで、宿主細胞に導入することによって、宿主細胞中でリボザイムを発現させることもできる。ベクターは当該宿主細胞中での遺伝子の発現を誘導可能なプロモーターを含む。

【0038】

本発明は、本発明のポリヌクレオチドの発現を抑制するsiRNA(small interfering RNA)を提供する。siRNAは、mRNAと同じ塩基配列を含む2本鎖のRNAが、当該mRNAに基づく蛋白質合成を阻害する作用(Fire et al. (1998) Nature 391: 806-811)を利用した、遺伝子発現制御技術である。2本鎖RNAによる遺伝子の発現抑制効果は、RNAi効果と呼ばれている。siRNAによる遺伝子発現制御は、マウスでも有効であったことが知られている(Zamore et al. (2000). Cell 101:25-33; Gur a (2000) Nature 404: 804-808)。つまり、2本鎖RNAの細胞への導入によって、遺伝子選択的に発現を抑制することができる。

【0039】

siRNAの長さは制限されない。細胞に導入された2本鎖のRNAは、細胞内で3'末端から21-23bpからなる断片に酵素的に切断される。2本鎖RNAを切断する酵素は、ダイサー(dicer)と呼ばれている。生成した2本鎖RNAの断片は、同じ配列を持つ標的塩基配列を認識して結合し、当該塩基配列がRNaseIII様のヌクレアーゼ活性によって切断される(Hammond et al. (2000) Nature, 404: 293-298; Zamore et al. (2000). Cell 101: 25-33)。

【0040】

RNAi効果による遺伝子の発現制御のためには、細胞内にsiRNAを導入する。siRNAはリボザイムと同様の方法により、細胞内に導入することができる。すなわち、化学的に合成された2本鎖RNAを細胞内に導入することができる。アンチセンスRNAであれsiRNAであれ、予め合成されたRNAを細胞に導入する場合には、ヌクレアーゼによる分解を防ぐために修飾しておくことができる。たとえば、チオ化されたRNAは、ヌクレアーゼの作用を受けにくい。

【0041】

あるいはsiRNAを細胞内で発現させることもできる。たとえば、センス配列とアンチセンス配列を挿入したベクターを宿主細胞に形質転換し、細胞内で発現さ

せることができる。センス鎖とアンチセンス鎖が連続して配置されていれば、ヘアピンループ構造を持つ2本鎖RNAが発現される。あるいは両者を異なるプロモーターの制御下に発現させて、別々のストランドからなる2本鎖RNAとすることもできる。siRNAの発現のためのプロモーターとしては、U6プロモーターなどが一般に用いられる。

【0042】

本発明のアンチセンスポリヌクレオチド、リボザイム、あるいはsiRNAを構成する塩基配列は、配列番号：1～配列番号：39に示した塩基配列と完全に同一、あるいは完全に相補的な塩基配列に加え、これらの塩基配列と高い相同性を有する塩基配列であっても良い。アンチセンスポリヌクレオチド、リボザイム、あるいはsiRNAの塩基配列において、高い相同性とは、通常90%以上、好ましくは95%以上、より好ましくは98%以上、更に好ましくは99%以上の相同性を言う。塩基配列の相同性は、たとえば本明細書に記載の方法によって明らかにすることができる。

【0043】

当業者は、発現を抑制すべき遺伝子の塩基配列に基づいて、siRNAをデザインすることができる。一般的なsiRNAのデザインのための方法として、たとえば以下の方法を示すことができる。すなわち、まずターゲット配列としては、次のような領域を避けるのが有利と考えられている。

5'側あるいは3'側の非翻訳領域

スタートコドン付近

これらの領域は、転写調節蛋白質の結合領域であることが多い。また多くのmRNAで保存された塩基配列を含む可能性があるので、目的以外の遺伝子に対して阻害作用を及ぼす場合がある。

したがって、たとえばスタートコドンよりも下流のORFの中にターゲット配列を設定するのが有利とされている。スタートコドンとターゲット配列との間隔は、たとえば50塩基以上とするのが望ましい。siRNAの塩基配列は、aaから始まり、19-21塩基の連続する塩基配列を含むようにするのが一般的である。またsiRNAの末端には2塩基のオーバーハングが付加される。オーバーハングの塩基配列

としては、dTdTやUUなどが用いられる。siRNAを構成する塩基配列のGC含量は、50%前後が好ましい。また一般にその分布は、siRNA全体で均一な方が好ましいとされている。

【0044】

siRNAの作用は、mRNAへの配列特異的なハイブリダイズに基づいている。したがって、できるだけ目的とする遺伝子に特異的な塩基配列をターゲット配列として選択することが、遺伝子に対して特異的な抑制効果を実現する上で重要な条件となる。したがって、ターゲット配列として選択した塩基配列は、相同性検索によって他の遺伝子との相同性が低いことを確認することが望ましい。塩基配列の相同性を決定するためのアルゴリズムは公知である。

【0045】

本発明のsiRNAは、本発明のポリヌクレオチドの発現を抑制する限り、上記のような一般的なデザイン方法によって得ることができる塩基配列に限定されない。たとえば、ターゲット配列が特定の遺伝子に特異的な塩基配列でなくても、相同な塩基配列を有する遺伝子が発現していない細胞においては、目的とする遺伝子に対する特異的な発現抑制作用を期待できる。更に、上記のような一般的なターゲット配列の選択方法によらなくても、RNAi効果を示す2本鎖RNAを得ることは可能である。

【0046】

本発明のDNA、あるいはそのアンチセンス、リボザイム、並びにsiRNAには、例えば、遺伝子治療への応用が考えられる。遺伝子治療の標的となる疾患としては、例えば、癌や各種炎症性疾患が好適であると考えられる。これら分子を遺伝子治療に用いる場合には、例えば、レトロウイルスベクター、アデノウイルスベクター、アデノ随伴ウイルスベクターなどのウイルスベクターやリポソームなどの非ウイルスベクターなどを利用して、ex vivo法やin vivo法などにより患者へ投与を行えばよい。

【0047】

また本発明は、本発明の蛋白質の部分ペプチドを含む。この部分ペプチドには、例えば、分泌蛋白質においてはシグナルペプチドが除去された蛋白質が含まれ

る。また、本発明の蛋白質が受容体やリガンドとしての活性を持つもの場合には、その競合阻害剤として機能する、受容体（あるいはリガンド）との結合能を有する部分ペプチドが含まれる。また、抗体調製のための抗原ペプチドが含まれる。部分ペプチドが本発明の蛋白質に特異的であるためには、少なくとも 7 アミノ酸、好ましくは 8 アミノ酸以上、より好ましくは 9 アミノ酸、更に好ましくは 10 アミノ酸以上のアミノ酸配列からなる。該部分ペプチドは、本発明の蛋白質に対する抗体や本発明の蛋白質の競合阻害剤の調製以外に、例えば、本発明の蛋白質に結合する蛋白質のスクリーニングなどに利用し得る。本発明の部分ペプチドは、例えば、遺伝子工学的手法、公知のペプチド合成法、あるいは本発明の蛋白質を適当なペプチダーゼで切断することによって製造することができる。

【 0 0 4 8 】

また、本発明は、本発明のポリヌクレオチドが挿入されたベクターに関する。本発明のベクターは、挿入した DNA を安定に保持するものであれば特に制限されない。例えば宿主に大腸菌を用いるのであれば、クローニング用ベクターとしては pBluescript ベクター (Stratagene 社製) などが好ましい。本発明の蛋白質を生産する目的においてベクターを用いる場合には、特に発現ベクターが有用である。発現ベクターは、試験管内、大腸菌内、培養細胞内、生物個体内で蛋白質を発現するベクターであれば特に制限されない。例えば、試験管内発現であれば pBES T ベクター (プロメガ社製)、大腸菌であれば pET ベクター (Invitrogen 社製)、培養細胞であれば pME18S-FL3 ベクター (GenBank Accession No. AB009864)、生物個体であれば pME18S ベクター (Mol Cell Biol. 8:466~472(1988)) などが好ましい。ベクターへの本発明の DNA の挿入は常法により制限酵素サイトを用いたりガーゼ反応により行うことができる (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.4 ~11.11)。

【 0 0 4 9 】

また、蛋白質発現を目的とした発現ベクターの構築に GATEWAY システム (インビトロジェン株式会社) という技術がある (実験医学 Vol. 18, No. 19 (12月号), p2716-2717, 2000)。このシステムはラムダファージのもつ 2 種類の部位特異

的組換え酵素（BPクロナーゼとLRクロナーゼ）とその特異的組換え部位をエントリーベクターと蛋白精製の際に有効なタグが既に挿入されているものも含むデスティネーションベクターにそれぞれ採用し、相同組換え反応を利用することによって発現ベクターを得るものである。

まず、1段階目の組換え反応を用いて目的のDNA断片をエントリーベクターへ挿入し、次にこの目的のDNA断片が挿入されたエントリーベクターとデスティネーションベクター間で2段階目の組換え反応をさせ、迅速かつ高効率で発現ベクターを得ることができる。前述のような制限酵素やリガーゼ反応を用いた定法では、発現ベクターを構築し目的の蛋白質を発現させるまで7～10日間程度の期日が必要となるが、GATEWAYではわずか3～4日間で目的の蛋白質の発現が可能になり、ハイスループットな発現蛋白質の機能解析が実現できる（<http://biotech.nikkeibp.co.jp/netlink/lto/gateway/>）。

【0050】

加えて本発明は、本発明のベクターを保持する形質転換体に関する。本発明のベクターが導入される宿主細胞としては特に制限はなく、目的に応じて種々の宿主細胞が用いられる。蛋白質を高発現させるための真核細胞としては、例えば、COS細胞、CHO細胞などを例示することができる。

宿主細胞へのベクター導入は、例えば、リン酸カルシウム沈殿法、電気パルス穿孔法（Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 9.1-9.9）、リポフェクタミン法（GIBCO-BRL社製）、マイクロインジェクション法などの方法で行うことが可能である。

【0051】

更に表1に示した配列番号に記載の塩基配列からなるポリヌクレオチド、またはその相補鎖に相補的な塩基配列からなる少なくとも1.5ヌクレオチドを含むポリヌクレオチドは、単に全長cDNAの合成のためのプライマーとして利用できるのみならず、全長cDNAによってコードされる本発明の蛋白質の異常を検査・診断するために利用できる。例えば、本発明のポリヌクレオチドをプライマーに用いたポリメラーゼ連鎖反応（ゲノムDNA-PCRやRT-PCR）により本発明の蛋白質をコー

ドするDNAを増幅することができる。また、全長cDNAの5'末端配列からゲノム配列上でのmRNA転写開始点が容易に特定可能なので、PCRやハイブリダイゼーションの手法を用いて5'上流の発現制御領域を容易に取得することができる。取得された遺伝子領域に対して、RFLP解析、SSCP、シーケンシング等の方法により、配列の異常を検査・診断することができる。特に本発明のmRNAの発現が特定の疾患によって変動する場合には、本発明のポリヌクレオチドをプローブやプライマーとして該mRNAの発現量を解析することによって該疾患の検出や診断を行うことができる。

【 0 0 5 2 】

また、本発明は、本発明の蛋白質に結合する抗体に関する。本発明の抗体の形態には特に制限はなく、ポリクローナル抗体やモノクローナル抗体または抗原結合性を有するそれらの一部も含まれる。また、全てのクラスの抗体が含まれる。さらに、本発明の抗体には、ヒト化抗体やキメラ抗体などの特殊抗体も含まれる。

本発明の抗体は、ポリクローナル抗体の場合には、常法に従いアミノ酸配列に相当するオリゴペプチドを合成して家兎に免疫することにより得ることが可能である (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.12~11.13)。一方、モノクローナル抗体は、常法に従い大腸菌で発現し精製した蛋白質を用いてマウスを免疫し、脾臓細胞と骨髓腫細胞を細胞融合させたハイブリドーマ細胞の中から得ることができる (Current protocols in Molecular Biology edit. Ausubel et al. (1987) Publish. John Wiley & Sons. Section 11.4~11.11)。

【 0 0 5 3 】

本発明の蛋白質に結合する抗体は、本発明の蛋白質の精製に加え、例えば、本発明の蛋白質の発現異常や構造異常の検査・診断に利用することも考えられる。具体的には、例えば組織、血液、または細胞などから蛋白質を抽出し、ウェスタンブロッティング、免疫沈降、ELISA等の方法による本発明の蛋白質の検出を通して、発現や構造の異常の有無を検査・診断することができる。

また、本発明の蛋白質に結合する抗体を、本発明の蛋白質に関連した疾患の治

療などの目的に利用することも考えられる。抗体を患者の治療目的で用いる場合には、ヒト抗体、ヒト化抗体、あるいはキメラ抗体が免疫原性の少ない点で好ましい。ヒト抗体は、免疫系をヒトのものと入れ換えたマウス（例えば、「Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice, Mendez, M.J. et al.(1997) Nat.Genet.15:146-156」参照）に免疫することにより調製することができる。また、ヒト化抗体は、モノクローナル抗体の超可変領域を用いた遺伝子組み換えによって調製することができる(Methods in Enzymology 203, 99-121(1991))。

【0054】

本発明のcDNA（クローン）がコードしているのは、例えば分泌・膜蛋白質、糖蛋白関連蛋白質、シグナル伝達関連蛋白質、転写関連蛋白質、疾患関連蛋白質、酵素・代謝関連蛋白質、細胞分裂・増殖関連蛋白質、細胞骨格関連蛋白質、核蛋白質・RNA合成関連蛋白質、蛋白質合成・輸送関連蛋白質、細胞防御関連蛋白質、発生・分化関連蛋白質、DNA・RNA結合蛋白質、ATP・GTP結合蛋白質のような機能が予測される蛋白質のアミノ酸配列である。これらの蛋白質としての機能を持つことは、相同性検索の結果から推定することができる。すなわち、本発明の全長cDNAの全長、もしくは部分塩基配列と相同性を示す公知の遺伝子・蛋白質を検索し、その遺伝子とそれがコードしている蛋白質の機能を参照すれば本発明のcDNAがコードしている蛋白質の機能を推定することができる。

【0055】

また、アミノ酸配列中にシグナル配列、膜貫通領域、核移行シグナル、糖鎖付加シグナル、リン酸化部位、及びZinc fingerモチーフ、SH3ドメイン等を見出すことでも本発明のcDNAがコードしている蛋白質の機能を推測できる。特にモチーフ、ドメインなどの構造はいくつかの蛋白質に共通して見出される部分配列構造で、蛋白質の最小限機能構造であり、現在までに機能が明らかとなっているもの、ないもの全て合わせてPfam(<http://www.sanger.ac.uk/Software/Pfam/index.shtml>)のVersion 7.7（2002年12月現在での最新版）においては4832種類が同定され、データベース化されている。

【0056】

具体的なモチーフ、ドメイン機能の一例として例えば、免疫反応に関与するT細胞において細胞膜上に発現するT Cell Receptorの細胞内領域に見い出されたITAM (immunoreceptor tyrosine-based activation motif) と呼ばれるモチーフ (Flaswinkel, H et.al. Semin Immunol 1995 Feb;7(1):21-7) はYXXL (チロシンー任意のアミノ酸ー任意のアミノ酸ーロイシン) がタンデムに並んだ構造をとっており、細胞外からの抗原や抗体刺激でこのモチーフの中のチロシン残基がリン酸化酵素ドメインを持った酵素 (LCK) によってリン酸化をうけ、そのリン酸化チロシンにZAP70がSH2ドメインを介して結合しシグナルが下流に伝わるということがわかっている (Bu, JY et.al. Proc Natl Acad Sci U S A 1995 May 23;92(11):5106-10, Neumeister, EN et.al. Mol Cell Biol 1995 Jun;15(6):3171-8)。

【 0 0 5 7 】

これに類似した現象はT細胞だけでなく、肥満細胞でも見い出されており (Chen, T et.al. J Biol Chem 1996 Oct 11;271(41):25308-15)、アレルギー、アトピー性皮膚炎、喘息などの免疫疾患において、分子レベルで見たときに免疫担当細胞が活性化する最初のプロセスとして理解されている。

【 0 0 5 8 】

上記は一例であるが、単純に記載したこれだけの反応においてもITAM、SH2ドメイン、蛋白質リン酸化酵素ドメインと3つの主要なモチーフ、ドメイン構造が主要な機能を担っており、しかもそれら3つの構造でこのメカニズムが説明できる。よってここで記載した免疫反応だけでなく、さまざまな細胞機能を分子メカニズムで理解するには、共通したモチーフ、ドメイン構造をもつ分子を集めカタログ化すること、そしてその最小構造の機能を解明すること、および未知蛋白質の機能解明にはまず第一にモチーフ、ドメイン構造の検索が非常に重要なことがわかる。また、蛋白質全体の構造はモチーフ、ドメインといった最小限構造の寄せ集めで成り立っており、その結果、タンパク質全体としての機能が発揮されると考えられている。

【 0 0 5 9 】

すなわち、ドメインやモチーフ構造の解析から、その蛋白質が全体として細胞内でどのような働きを担っているかということを分子レベルで極めて正確に予測

することが可能である。また、一部アミノ酸配列とGFP蛋白質などとの融合タンパクを作製して、培養細胞等に導入し、例えば細胞膜に局在すれば受容体やイオンチャンネルといった機能を持つ可能性が示唆されたり、核に局在すればDNAに結合したり転写に関与するといった機能予測が可能である。このように蛋白質の局在を調べることでも類推できる。

【0060】

本発明で得られた全長cDNAはその全塩基配列、およびそれがコードするアミノ酸配列をもとに上記のような解析を行うことで機能予測が可能であるが、cDNAの配列が全塩基配列でなくても部分的な配列情報(好ましくは300塩基以上)があれば機能予測は可能であることが多い。しかし、部分的な配列情報をもとにした相同性検索からの機能予測は、必ずしも全塩基配列をもとに予測された機能と一致しない場合があり、全塩基配列をもとにした機能予測のほうが好ましいのはいうまでもない。

【0061】

機能予測のより具体的方法として、相同性検索の場合はGenBank、Swiss-Prot、UniGene、nr、RefSeqといった各データベースを対象にBLASTやFASTAなどで相同性検索を行い、ヒットした遺伝子とそれがコードしている蛋白質の機能を参照することで本発明のcDNAがコードしている蛋白質の機能を推定する。また、構造からの予測においては全塩基配列から推定されたアミノ酸配列に対して、シグナル配列、膜貫通領域の予測ならばPSORT [K. Nakai & M. Kanehisa, Genomics, 14: 897-911 (1992)]やSOSUI [T.Hirokawa et.al. Bioinformatics, 14, 378-379 (1998)] (三井情報開発株式会社販売)、MEMSAT [D.T.Jones, W.R.Taylor & J.M.Thornton, Biochemistry, 33, 3038-3049 (1994)] など、またモチーフやドメインの予測ならばPfamやPROSITE (<http://www.expasy.ch/prosite/>)等に対して検索を行うことによって、クローン中にコードされる蛋白質のより詳細な機能予測が可能である。

【0062】

このようにして、本発明の全塩基配列が明らかになった全長39クローンについて、GenBank、Swiss-Prot、nr、RefSeq の各データベースを対象に相同性検索

を行った（実施例 4 および相同性検索結果参照）。また全長塩基配列から推定されたアミノ酸配列に対してPSORT、SOSUIを用いたシグナル配列、および膜貫通領域の検索を行った（実施例 5 参照）。これらの結果から、アノテーションを基本とした機能予測（Swiss-Protのヒットデータであればキーワードを参照する。nr、RefSeqのヒットデータであればDefinitionやReference情報を参照する）、および推定ORFに対するPSORTを用いたシグナルシーケンス検索、SOSUIを用いた膜貫通領域の検索結果をあわせて、以下 1 3 種類の機能カテゴリーへの分類を行った。結果、3 3 クローンについては以下のカテゴリーに属する蛋白質をコードしていることが推定された。

【0 0 6 3】

分泌・膜蛋白質（1 9 クローン）
糖蛋白関連蛋白質（1 0 クローン）
シグナル伝達関連蛋白質（4 クローン）
転写関連蛋白質（7 クローン）
疾患関連蛋白質（2 3 クローン）
酵素・代謝関連蛋白質（8 クローン）
細胞分裂・増殖関連蛋白質（2 クローン）
細胞骨格関連蛋白質（1 クローン）
核蛋白質・RNA合成関連蛋白質（1 クローン）
蛋白質合成・輸送関連蛋白質（2 クローン）
発生・分化関連蛋白質（1 クローン）
DNA・RNA結合蛋白質（8 クローン）
ATP・GTP結合蛋白質（2 クローン）

【0 0 6 4】

分泌・膜蛋白質に属すると推定されたクローンは、以下の 1 9 クローンであった。

ADRGL2011190, BRACE3046450, BRAWH3013197, BRAWH3028645, BRAWH3046240, BRTHA3024233, PROST2010326, TBAES2004105, TBAES2007379, TBAES2007481, TBAES2008133, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3044175, THYMU3

046350, TLIVE2007192, UTERU2025415, UTERU2036507

【 0 0 6 5 】

糖蛋白質関連蛋白質に属すると推定されたクローンは、以下の 1 0 クローンであった。

BRAWH3013197, BRAWH3046240, PROST2010326, TBAES2004105, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3046350, TRACH3021066, UTERU2036507

【 0 0 6 6 】

シグナル伝達関連蛋白質に属すると推定されたクローンは、以下の 4 クローンであった。

BRHIP3030064, CTONG2003764, PLACE7013963, TRACH3021066

【 0 0 6 7 】

転写関連蛋白質に属すると推定されたクローンは、以下の 7 クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 0 6 8 】

疾患関連蛋白質に属すると推定されたクローンは、以下の 2 3 クローンであった。

ADRGL2011190, BRACE3002184, BRACE3026993, BRACE3046450, BRAWH3013197, BRAWH3046240, BRCAN2019772, BRHIP3030064, BRTHA3024233, CTONG2002832, CTONG2003764, PROST2010326, TBAES2004105, TESTI2046536, TESTI4005158, TESTI4005500, THYMU3024602, THYMU3046350, TRACH1000193, TRACH3019290, TRACH3021066, UTERU2017492, UTERU2036507

【 0 0 6 9 】

また上記 2 3 クローンは全てが、Swiss-Prot ヒットデータ、及び nr、RefSeq ヒットデータが、ヒトの遺伝子と疾患のデータベースである Online Mendelian Inheritance in Man (OMIM) に登録されている遺伝子、蛋白質であった。以下にクローン名と OMIM Number を示す (クローン名の後ろのカッコ内は対象となった OMIM Number)。

ADRGL2011190 (602658), BRACE3002184 (603899), BRACE3026993 (601282), BRACE3046450 (603035), BRAWH3013197 (604100), BRAWH3046240 (602646), BRCAN2019772 (603849), BRHIP3030064 (605762), BRTHA3024233 (603088), CTONG2002832 (604668), CTONG2003764 (176977), PROST2010326 (604110), TBAES2004105 (605008), TESTI2046536 (182890; 182882), TESTI4005158 (601328), TESTI4005500 (602277), THYMU3024602 (600493), THYMU3046350 (603080), TRACH1000193 (139139), TRACH3019290 (603975), TRACH3021066 (164870), UTERU2017492 (602917), UTERU2036507 (605008)

【 0 0 7 0 】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 8 クローンであった。

ADRGL2011190, BRHIP3030064, CTONG2003764, PLACE7013963, TBAES2004105, THYMU3044175, TRACH3021066, UTERU2036507

【 0 0 7 1 】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

TBAES2007481, TESTI2046536

【 0 0 7 2 】

細胞骨格関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

BRACE3026993

【 0 0 7 3 】

核蛋白質・RNA合成関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

TESTI2046536

【 0 0 7 4 】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

TBAES2004105, UTERU2036507

【 0 0 7 5 】

発生、分化関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

BRCAN2019772

【 0 0 7 6 】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 8 クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI2046536, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 0 7 7 】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

CTONG2003764, TRACH3021066

【 0 0 7 8 】

以下の 6 クローンについては、現在のところ相同性検索の情報からは機能を推定できる情報の得られないクローンであった。これらクローンについては今後、データベースのアップデートによって機能が明らかになる可能性がある。クローン名を以下に示した。

BRAMY3008096, BRAMY3016953, BRHIP3038037, BRTHA3004432, TESTI4052089, UTERU3020583

【 0 0 7 9 】

なお、蛋白質の機能が必ずしも上記に示す機能カテゴリーの一つのみに属するわけではないため、いずれで予測された機能カテゴリーにも該当する可能性がある。またこれらの機能カテゴリーで分類されたクローンには、今後の解析により新たな機能が付加される可能性がある。

【 0 0 8 0 】

また、本発明の全塩基配列が明らかになった全長 3 9 クローンについて、推定されたアミノ酸配列のPfam (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>) に対するドメイン検索の結果（実施例 5 参照）から得られるヒットデータ

のドメイン、モチーフ名やアクセッション番号を用いて、Pfamのサイト内やInterPro (<http://www.ebi.ac.uk/interpro/>)、PROSITE (<http://www.expasy.ch/cgi-bin/prosite-list.pl>)等の各リンク先における各ドメイン、モチーフの詳細な説明や、特にPROSITEにおいては独自の機能カテゴリー分類を参照することができる。このようにして、Pfamでヒットした29クローン中にコードされるタンパク質の機能予測を行い、以下9種類の機能カテゴリーへの分類を行った。結果、22クローンについては以下のカテゴリーに属する蛋白質をコードしていることが推定された。

【0081】

分泌・膜蛋白質 (10クローン)
糖蛋白質関連蛋白質 (2クローン)
シグナル伝達関連蛋白質 (3クローン)
転写関連蛋白質 (8クローン)
酵素・代謝関連蛋白質 (7クローン)
細胞分裂・増殖関連蛋白質 (1クローン)
蛋白質合成・輸送関連蛋白質 (1クローン)
DNA・RNA結合蛋白質 (7クローン)
ATP・GTP結合蛋白質 (1クローン)

【0082】

分泌・膜蛋白質に属すると推定されたクローンは、以下の10クローンであった。

BRAWH3013197, BRAWH3028645, BRAWH3046240, BRCAN2019772, PROST2010326, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3046350, TRACH1000193

【0083】

糖蛋白質関連蛋白質に属すると推定されたクローンは、以下の2クローンであった。

BRAWH3028645, TESTI4005158

【0084】

シグナル伝達関連蛋白質に属すると推定されたクローンは、以下の3クローン

であった。

BRAMY3008096, TESTI4052089, TRACH3021066

【 0 0 8 5 】

転写関連蛋白質に属すると推定されたクローンは、以下の 8 クローンであった

。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI2046536, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 0 8 6 】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 7 クローンであった。

ADRGL2011190, BRAMY3008096, CTONG2002832, TESTI4052089, THYMU3024602, THYMU3044175, TRACH3021066

【 0 0 8 7 】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

TBAES2007481

【 0 0 8 8 】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

CTONG2002832

【 0 0 8 9 】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 7 クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 0 9 0 】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

ADRGL2011190

【0091】

以下の7クローンについては、Pfamでヒットデータ（実施例5参照）があったものの、上記のいずれのカテゴリに属するか明らかでないクローンであった。今後同様のドメイン、モチーフを持つタンパク質のデータの蓄積と共に機能がより詳細に解明され、上記のカテゴリに分類できる可能性がある。

BRACE3026993// TSC-22/dip/bun family

BRACE3046450// PDZ domain (Also known as DHR or GLGF). //PDZ domain (Also known as DHR or GLGF). //PDZ domain (Also known as DHR or GLGF). 1//PDZ domain (Also known as DHR or GLGF).

CTONG2003764// Phorbol esters/diacylglycerol binding domain (C1 domain)

TBAES2004105// Thrombospondin type 1 domain

TBAES2007379// EGF-like domain//EGF-like domain//Trypsin Inhibitor like cysteine rich domain//EGF-like domain//EGF-like domain//Keratin, high sulfur B2 protein//EGF-like domain//EGF-like domain//Granulins//EGF-like domain//EGF-like domain//EGF-like domain

TBAES2008133// PDZ domain (Also known as DHR or GLGF).

UTERU2036507// Thrombospondin type 1 domain//NTR/C345C module

【0092】

なお、モチーフやドメインの機能が必ずしも上記に示す機能カテゴリの一つのみに属するわけではないため、いずれで予測された機能カテゴリにも該当する可能性がある。またこれら以外にPfamでヒットデータがなかった残りのクローンについても、今後タンパク質のデータの蓄積と共に新たなドメイン、モチーフが見い出された場合、再びクローンの推定アミノ酸配列を新しいデータベースに対して解析することで新たな機能を有したドメイン、モチーフが発見され、カテゴリ分類できる可能性がある。

【0093】

これらクローンにコードされる蛋白質は、いずれも全長アミノ酸配列を備えることから、適当な発現系を適用して組み換え体として発現させたり、細胞にインジェクションすることにより、あるいは、そのタンパクを特異的に認識する抗体

を作製し、用いることで、その生物学的活性、及び細胞増殖・分化といった細胞状態変化への作用を解析することが可能である。

【 0 0 9 4 】

各蛋白質は、それぞれ次に示すような手法にもとづいて、それぞれの蛋白質の生物学的活性の解析が可能である。

分泌蛋白質、膜蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Ion Channels』(R.H. Ashley編、1995)、

『Growth Factors』(I. McKay, I. Leigh編、1993)、『Extracellular Matrix』(M.A. Haralson, J.R. Hassell編、1995)

糖蛋白質関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Glycobiology』(M. Fukuda, A. Kobata編、1993)、

「Method in Molecular Biology」(Humana Press社)シリーズの『Glycoprotein Analysis in Biomedicine』(Elizabeth F. Hounsell編、1993)、

シグナル伝達関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Signal Transduction』(G. Milligan編、1992)、

『Protein Phosphorylation』(D.G. Hardie編、1993)、または「Method in Molecular Biology」(Humana Press社)シリーズの『Signal Transduction Protocols』(David A. Kendall, Stephen J. Hill編、1995)、

転写関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Gene Transcription』(B.D. Hames, S.J. Higgins編、1993)、

『Transcription Factors』(D.S. Latchman編、1993)、

酵素・代謝関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Enzyme Assays』(ROBERT EISENTHAL and MICHAEL J. DANSON編、1992)、

細胞分裂・増殖関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Cell Growth, Differentiation and Senescence』(GEORGE STUDZINSKI編、2000)、
細胞骨格関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Cytoskeleton: Signaling and Cell Regulation』(KERMIT L. CARRAWAY and CAROLIE A. CAROTHERS CARRAWAY編、2000)、

「Method in Molecular Biology」(Humana Press社)シリーズの『Cytoskeleton Methods and Protocols』(Gavin; Ray H. 編、2000)、
核蛋白質・RNA合成関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Nuclear Receptors』(DIDIER PICARD編、1999)、

『RNA Processing』(STEPHEN J. HIGGINS and B. DAVID HAMES編、1994)、
蛋白質合成・輸送関連蛋白質：

「The Practical Approach Series」(IRL PRESS社)の『Membrane Transport』(STEPHEN A. BALDWIN編、2000)、

「Method in Molecular Biology」(Humana Press社)シリーズの『Protein Synthesis Methods and Protocols』(Martin, Robin編、1998)、
細胞防御関連蛋白質：

「Method in Molecular Biology」(Humana Press社)シリーズの『DNA Repair Protocols』(Henderson, Daryl S.、1999)、

『Chaperonin Protocols』(Schneider, Christine編、2000)、
発生・分化関連蛋白質：

「Method in Molecular Biology」(Humana Press社)シリーズの『Developmental Biology Protocols』(ROBERT EISENTHAL and MICHAEL J. DANSON編、1992)、
DNA・RNA結合蛋白質：

「Method in Molecular Biology」(Humana Press社)シリーズの『DNA-Protein Interactions Principles and Protocols』(Kneale, G. Geoff編、1994)、

『RNA-Protein Interaction Protocols』(Haynes, Susan R. 編、1999)、
ATP・GTP結合蛋白質：

「Method in Molecular Biology」(Humana Press社)シリーズの『Signal Transduction Protocols』(David A. Kendall, Stephen J. Hill編、1995)

これら以外の手法については、Methods in Enzymology(Academic Press)を参照して蛋白質の活性を解析することができる。

【 0 0 9 5 】

なお、上述したカテゴリー分類において、分泌・膜蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中に growth factor, cytokine, hormone, signal, transmembrane, membrane, extracellular matrix, receptor, G-protein coupled receptor, ionic channel, voltage-gated channel, calcium channel, cell adhesion, collagen, connective tissue 等、分泌・膜蛋白質と推定される記載があった、もしくはPSORTとSOSUIによる推定ORFの解析の結果、シグナルシーケンスや膜貫通領域があった、またPfamによるドメイン、モチーフ検索の結果、受容体、イオンチャネル、ホルモン、成長因子などと推測されるような例えば7 transmembrane receptor, Pancreatic hormone peptides, Ion transport protein, Fibroblast growth factor等のドメイン、モチーフがあったクローンである。

【 0 0 9 6 】

糖蛋白質関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中に glycoprotein 等、糖蛋白質関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、糖タンパク質、糖転移酵素などGlycobiologyに関わると推測されるような例えばImmunoglobulin domain, Glycosyl transferases group 1等のドメイン、モチーフがあったクローンである。

【 0 0 9 7 】

シグナル伝達関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中に serine/threonine-protein kinase, tyrosine-protein kinase, SH3 domain, SH2 domain等、シグナル伝達関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、タンパク質リン酸化酵素、脱リン酸化酵素、SH2ドメイン、Small Gタンパク質など

と推測されるような例えばEukaryotic protein kinase domain, Protein phosphatase 2C, Ras family等のドメイン、モチーフがあったクローンである。

【 0 0 9 8 】

転写関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中に transcription regulation, zinc finger, homeobox 等、転写関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、転写因子や転写調節に関わるタンパク質などと推測されるような例えばbZIP transcription factor, Zinc finger, C2H2 type等のドメイン、モチーフがあったクローンである。

【 0 0 9 9 】

疾患関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中に disease mutation, syndrome 等、疾患関連蛋白質と推定される記載があった、あるいは全長塩基配列に対するSwiss-Prot、GenBank、UniGene、nr、RefSeqヒットデータが、後述するヒトの遺伝子と疾患のデータベースであるOnline Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>)に登録されている遺伝子や蛋白質であった、また、Pfamによるドメイン、モチーフ検索の結果、特定の疾患で発現が見られるようなタンパク質や、疾患で発現が上昇したり減少したりすると推測されるような例えばWilm's tumour protein, von Hippel-Lindau disease tumor suppressor protein等のドメイン、モチーフがあったクローンである。

【 0 1 0 0 】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にmetabolism, oxidoreductase, E.C.No. (Enzyme commission number)等、酵素・代謝関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、転移酵素、合成酵素、加水分解酵素などと推測されるような例えばAldehyde dehydrogenase family, Chitin synthase, Glucose-6-phosphate dehydrogenase等のドメイン、モチーフがあったクローンである。

【 0 1 0 1 】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、cell division, cell cycle, mitosis, chromosomal protein, cell growth, apoptosis等、細胞分裂・増殖関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、サイクリン、細胞増殖制御タンパク質などと推測されるような例えばCyclin, Cell division protein等のドメイン、モチーフがあったクローンである。

【0102】

細胞骨格関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にstructural protein, cytoskeleton, actin-binding, microtubules等、細胞骨格関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、アクチン、キネシン、フィブロネクチンなどと推測されるような例えばActin, Fibronectin type I domain, Kinesin motor domain等のドメイン、モチーフがあったクローンである。

【0103】

核蛋白質・RNA合成関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にnuclear protein, RNA splicing, RNA processing, RNA helicase, polyadenylation等、核蛋白質・RNA合成関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、スプライシング因子、RNA合成酵素、ヘリカーゼなどと推測されるような例えばHepatitis C virus RNA dependent RNA polymerase, DEAD/DEAH box helicase等のドメイン、モチーフがあったクローンである。

【0104】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にtranslation regulation, protein biosynthesis, amino-acid biosynthesis, ribosomal protein, protein transport, signal recognition particle等、蛋白質合成・輸送関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、翻訳関連タンパク質、ユビキチン関連タンパク質、Ribosomal proteinなどと推測されるような例えばTranslation initiation factor SUI1, Ubiquitin family, Ribosomal

protein Ll6等のドメイン、モチーフがあったクローンである。

【0105】

細胞防御関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にheat shock, DNA repair, DNA damage等、細胞防御関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、分子シャペロン、DNA修復タンパク質などと推測されるような例えばHsp90 protein, DNA mismatch repair protein等のドメイン、モチーフがあったクローンである。

【0106】

発生・分化関連蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にdevelopmental protein等、発生・分化関連蛋白質と推定される記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、器官形成関連タンパク質などと推測されるような例えばFloricaula / Leafy protein等のドメイン、モチーフがあったクローンである。

【0107】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にDNA-binding, RNA-binding等と記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、転写因子、DNAリガーゼをはじめとしたDNA・RNA関連酵素類、Zinc-finger関連タンパク質などと推測されるような例えばTranscription factor WhiB, B-box zinc finger, tRNA synthetases class I (C)等のドメイン、モチーフがあったクローンである。

【0108】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンとは、相同性検索の結果、ヒットデータ中にATP-binding, GTP-binding等と記載があった、あるいはPfamによるドメイン、モチーフ検索の結果、ATPase等をはじめとしたATP・GTP関連酵素類、Gタンパク質などと推測されるような例えばE1-E2 ATPase, Ras family等のドメイン、モチーフがあったクローンである。

【0109】

疾患関連蛋白質については、前述したように機能ごとの解析が可能であるほか

、疾患関連蛋白質を発現して得られた特異認識抗体を用いて、特定の疾患と蛋白質の発現量や活性との相関を知ることができる。あるいは、ヒトの遺伝子と疾患のデータベースであるOMIMを利用し、解析が可能である。なおOMIMには常に新しい情報が付加されている。したがって当業者は、特定の疾患と本発明の遺伝子との新たな関係を最新のデータベースから見出すことができる。疾患関連蛋白質は、診断マーカー、発現・活性の増減を制御する薬剤、あるいは遺伝子治療のターゲットになるなど医薬品の開発等に有用である。

【0 1 1 0】

また、分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質をはじめ、上記の14種類のカテゴリの蛋白質に限らず、種々の機能をもつ蛋白質についても、OMIMを利用してキーワードで検索すると、各キーワードにおいて、多くの疾患に関連した結果が得られた（分泌、膜蛋白質について、OMIMで検索した結果を一例として以下に示す）。あるいは、例えば転写関連蛋白質やシグナル伝達関連蛋白質については、疾患との関連がそれぞれ、藤井・田村・諸橋・影山・佐竹編の実験医学増刊「転写因子研究1999」Vol.17, No.3, (1999)や、遺伝子医学Vol.3, No.2(1999)で報告されている。例えば、がんを例に挙げると、裳華房生命科学シリーズ「がんの生物学」（松原聡著、1992）にあるように、がんには分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質ばかりでなく、酵素・代謝関連蛋白質、細胞骨格関連蛋白質、細胞分裂・増殖関連蛋白質といった多くの蛋白質が関与することが示されている。このように、疾患関連蛋白質ばかりでなく、分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質等も疾患に関与することが多く、医療産業上のターゲットとして、有用なことがわかる。

【0 1 1 1】

一例として、分泌、膜蛋白質について、OMIMで検索した結果を以下に示す。OMIM検索に用いたキーワードには、

- (1) secretion protein
- (2) membrane protein
- (3) channel

(4) extracellular matrix

を用いた。

【 0 1 1 2 】

検索結果には、OMIM登録番号記号のみを記載した。なお、番号は検索結果表示時に最初に現れる 5 0 件のみ記した。この番号をもとにOMIMで疾患と遺伝子や蛋白質との関係を示すデータを見ることができる。また、OMIMデータは日々更新されている。

1) Secretion protein (分泌蛋白質)

このキーワードで検索された疾患と関連のある遺伝子は 4 3 6 登録されており、そのうちの 5 0 遺伝子は以下のOMIM番号をもつ。

*604667、*104760、*176860、*139320、*118910、*151675、*107400、*604029、
#200100、*177061、*600946、*601693、*139250、*176880、*600998、*603850、
*605083、*147572、*179513、*606055、*604028、*125950、*157147、*246700、
*602926、*600560、*602421、*603215、185860、*600174、*179512、*109270、*
179511、*179510、*179509、*601146、*604710、*177020、*138120、*170280、*
600626、*164160、*168470、*154545、*603831、*601652、*104311、*601489、*
603062、*102720

【 0 1 1 3 】

2) Membrane protein (膜蛋白質)

このキーワードで検索された疾患と関連のある遺伝子は 1 8 7 3 登録されており、そのうちの 5 0 遺伝子は以下のOMIM番号をもつ。

*130500、*605704、*305360、*153330、*109270、*173610、*170995、*120920、
*170993、*309060、*104776、*602333、*605703、*602690、*605943、*159430、
*600897、*606867、*133090、*601178、*602413、*602003、*604405、*605940、
*603237、*109280、*606958、*600378、*606959、*602173、*107776、*602334、
*125305、*602335、*309845、*601134、*605731、*606795、*185881、*607178、
*603177、*154045、*603214、*603718、*606909、*600594、*603241、*606629、
*603657、*600182

【 0 1 1 4 】

3) Channel (膜蛋白質のメンバー)

このキーワードで検索された疾患と関連のある遺伝子は 4 4 9 登録されており、そのうちの 5 0 遺伝子は以下のOMIM番号をもつ。

*176266、*600724、*605427、*182390、*123825、*114208、*114206、*114205、
*176267、*600053、*601784、*603749、*182392、*600937、*603415、*114204、
*114209、*114207、*607370、*604528、*604527、*601011、*600760、*192500、
*118425、*600228、*600359、*176261、*602235、*600761、*182389、*300008、
*600877、*605692、*300338、*602232、*603537、*182391、*176263、*602343、
*601328、*605874、*604385、*603939、*602208、*601534、*601958、*603220、
*600504、*607368

【0 1 1 5】

4) Extracellular matrix

このキーワードで検索された疾患と関連のある遺伝子は 2 6 7 登録されており、そのうちの 5 0 遺伝子は以下のOMIM番号をもつ。

*605912、*602201、*603479、*604633、*601418、*601548、*115437、*154870、
*120361、*602285、*600754、*602262、*134797、*602261、*603320、*603321、
*604871、*604629、*601807、*154700、*128239、*600310、*605470、*185250、
*178990、*603767、*120360、*185261、*116935、*607056、*253700、*190180、
*600985、*188826、*193300、*276901、*308700、*120150、*602109、*120324、
*600514、*177170、*247100、*116920、*200610、*605127、*601313、*601652、
*120180、*154790

また、これらと同様に、前述のカテゴリー分類のところに示した各種のキーワード等もOMIMの検索に用いることによって疾患との関連をみることができる。

【0 1 1 6】

また、本発明のcDNAの塩基配列を用いれば、そのcDNAの塩基配列を有する遺伝子の発現頻度を解析することができる。更にこうして解析された発現頻度情報に基づいて、当該遺伝子の機能を予測することができる。

【0 1 1 7】

疾患に関連した遺伝子を調べる方法として病態組織と正常組織において遺伝子

発現量の違いを調べる発現頻度解析がある。発現頻度解析には、ノーザンブロットティング法やRT-PCR法、およびDNAマクロアレイやDNAマイクロアレイを用いた発現頻度解析法がある(実験医学 Vol.17, No.8, 980-1056 (1999)、村松・那波監修 細胞工学別冊「DNAマイクロアレイと最新PCR法」(秀潤社, 2000))。更に、こういった解析方法以外に、発現している遺伝子の塩基配列をコンピューターを利用した解析で比較することによっても発現頻度を解析することができる。例えば、BODYMAPと呼ばれるデータベースは、様々な組織・細胞のcDNAライブラリーから、無作為に遺伝子クローンを抽出し、3'末端領域の塩基配列の相同性情報をもとにして、相同性のあるものはまとめてクラスターとすることによって、クラスター単位で遺伝子を分類して、各クラスターに含有されるクローンの個数を比較することによって遺伝子の発現頻度情報を得ている (<http://bodymap.ims.u-tokyo.ac.jp/>)。

【0118】

このような解析手法により、病態組織と正常組織において遺伝子発現量の違いを調べた結果から発現量の違いが明らかな遺伝子は、その疾患に関連した遺伝子といえる。また、病態組織でなくとも、病態に関連した特異的な現象を再現させた培養細胞と正常細胞において遺伝子発現量の違いを調べた結果から発現量の違いが明らかな遺伝子は、その疾患に関連した遺伝子といえる。

【0119】

全塩基配列が明らかになった39クローンについて、以下のデータベースを用いて、特定の病態や機能に関連する遺伝子を選択した(実施例8. 「in silicoにおける発現頻度解析」参照)。本発明の解析に用いたデータベースは、1,402,070個のクローンの塩基配列をデータベース化したものであり、解析母数としては十分なデータベースである。このデータベースを構成している配列情報は、実施例1に示した様々な組織や細胞由来のcDNAライブラリーからcDNAクローンを無作為に選択して、その5'末端領域の配列を決定することによって得た。

【0120】

次にこのデータベースにある各クローンの塩基配列を、塩基配列の相同性検索プログラムによって相同な配列同士をカテゴライズし(クラスター化)、各クラ

スターに属するクローン数を各ライブラリー毎に集計し規格化することによって、ある遺伝子のcDNAライブラリー内での存在比を解析した。この解析によって、cDNAライブラリーのソースとなっている組織や細胞における、ある遺伝子の発現頻度情報を得た。

【0 1 2 1】

次に本発明のcDNAの塩基配列を持つ遺伝子の、組織や細胞間での発現を解析するために、大量のcDNAクローンを解析した組織や細胞由来のライブラリーを組織・細胞間での発現量の比較の対象にした。すなわち 6 0 0 個以上のcDNAクローンの塩基配列を解析した組織や細胞について、先に規格化した数値を組織間や細胞間で比較し、遺伝子の発現頻度の変化を解析した。この解析によって以下に続く病態や機能に関連する遺伝子であることが示された。なお、以降に示される表 3 ～表 9、表 1 0、および表 1 1 ～表 1 3 中の各数値は、相対的な発現頻度を示し、数値が大きいほど発現量が多いことを示す。なお、表の一部に比較したライブラリー間ではさほど大きな差がない遺伝子も含まれるが、他の組織、細胞由来のライブラリーと比較した場合は、有意な差が認められるので、それぞれの組織、細胞に特異的な遺伝子であり、疾患の診断マーカーや、分子メカニズム解明に有用な遺伝子と言える。

【0 1 2 2】

アルツハイマー病関連遺伝子

アルツハイマー病とは記憶力が低下し、進行すれば生活が困難となり介護が必要となる脳神経系の疾患であり、進行すれば脳そのものが萎縮する。その発症の要因はストレスなどの環境因子、高血圧やコレステロール血症などの血管因子も関わりがあるといわれているが、未だ不明である。したがって、正常脳組織とアルツハイマーの病態組織を比較した時、発現に差のある遺伝子はアルツハイマー病に関連する遺伝子であり、病態の発症メカニズムの解明や、遺伝子診断に有用であると考えられる。アルツハイマー患者の大脳皮質由来のライブラリー (BRALZ、BRASW) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 3)、両者で発現変化のある遺伝子は以下の 5 クローンであった。

。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子はアルツハイマー病に関する遺伝子である。

【 0 1 2 3 】

パーキンソン病関連遺伝子

パーキンソン病とは脳内の黒質で作られるドーパミンという神経伝達物質が十分量作られなくなり、その結果、手が震え、筋肉の動きが固くなって身体の動きが鈍くなる等の運動障害を引き起こす脳神経系の疾患である。脳の神経細胞は通常、歳を取るにつれて少しずつ減少するが、パーキンソン病では黒質の神経細胞が普通よりも早く著しく減少する。よって脳組織全体と黒質とを比較した時、発現に差のある遺伝子は黒質特異的な変動をするパーキンソン病に関連する遺伝子であり、発症メカニズムの解明や遺伝子診断に有用であると考えられる。黒質由来のライブラリー (BRSSN) と、正常全脳組織由来のライブラリー (BRAWH) の cDNA を解析して比較した結果 (表 4)、両者で発現変化のある遺伝子は以下の 5 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子はパーキンソン病に関する遺伝子である。

【 0 1 2 4 】

短期記憶・痴呆症に関連する遺伝子

脳組織の中で海馬とは記憶を扱う非常に重要な部位であり、得た情報の情報の要・不要を判断して、他の脳部位に記憶を蓄えさせる、記憶固定の働きがある。臨床的知見より、海馬に異常をきたしたり最悪海馬が無くなると、5分程度しか新しいことを覚えていられなくなる。また痴呆症患者の一部はこの海馬に異常をきたしていると考えられている。脳組織全体と海馬とを比較した時、発現に差のある遺伝子は記憶に関与したり、痴呆症に関連する遺伝子であり、記憶のメカニズム解明や遺伝子診断に有用であると考えられる。海馬由来のライブラリー (BRHIP) と、正常全脳組織由来のライブラリー (BRAWH) の cDNA を解析して比較した結果 (表 5)、両者で発現変化のある遺伝子は以下の 5 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子は記憶および痴呆症に関する遺伝子である。

【0 1 2 5】

平衡感覚・運動機能に関する遺伝子

小脳は平衡感覚と筋肉運動、運動学習の中枢である。この領域は運動の調節に関与していると考えられており、小脳が動作することによって無意識的にスムーズな運動をすることが可能になる。また、運動だけでなく読み書きなどより高次の運動の慣れにも小脳が関与していることも最近の研究で解明されつつある。脳組織全体と小脳とを比較した時、発現に差のある遺伝子は平衡感覚や運動機能に関与する遺伝子であり、脳が制御する運動機能の分子メカニズム解明に有用であると考えられる。小脳由来のライブラリー (BRACE) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 6)、両者で発現変化のある遺伝子は以下の 9 クローンであった。

BRACE3002184, BRACE3026993, BRACE3046450, BRAMY3008096, BRAWH3013197, BRAWH3028645, BRTHA3004432, TESTI4002988, TRACH1000193

これらの遺伝子は平衡感覚および運動機能に関する遺伝子である。

【0 1 2 6】

感覚器からの情報伝達に関する遺伝子

視床は、大脳と結びつきの強い神経細胞が集まった部分であり、脊髄などから伝わってきた感覚情報を大脳の関係部分に伝えたり、大脳の運動の指令を調節する。例えば視覚では映像を大きさ、形、色に分け、聴覚では音声を音量、耳障りの良し悪しで分け、大脳皮質の感覚野に送る。脳組織全体と視床とを比較した時、発現に差のある遺伝子は感覚器からの情報伝達に関与する遺伝子であり、脳が制御する情報伝達の分子メカニズム解明に有用であると考えられる。視床由来のライブラリー (BRTHA) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 7)、両者で発現変化のある遺伝子は以下の 8 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, BRTHA3024233, CTONG2002832, THYMU3046350, TRACH1000193

これらの遺伝子は感覚器からの情報伝達に関する遺伝子である。

【0 1 2 7】

情動反応に関する遺伝子

扁桃は脳の感情中枢である。扁桃を通過した情報は感情反応、例えばパニックや恐怖反応などを引き起こす。刺激が扁桃で情動評価されて強い恐怖を生じたとき、扁桃は脳の各部に警戒信号を出す。その結果、手の平の発汗、心悸亢進、血圧上昇、アドレナリンの急激分泌等の反応が起きる。いわば扁桃体は身体に警戒信号を送り、その結果として体を警戒態勢に入らせる一種の防衛本能を司っている組織とも言える。脳組織全体と扁桃とを比較した時、発現に差のある遺伝子は情動反応に関与する遺伝子であり、感情反応や恐怖反応、パニックなどの分子メカニズム解明に有用であると考えられる。扁桃由来のライブラリー (BRAMY) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 8)、両者で発現変化のある遺伝子は以下の 9 クローンであった。

BRACE3046450, BRAMY3008096, BRAWH3013197, BRAWH3028645, BRTHA3004432, TESTI4002988, THYMU3046350, TRACH1000193, TRACH3019290

これらの遺伝子は情動反応に関する遺伝子である。

【 0 1 2 8 】

癌関連遺伝子

癌の組織では、正常組織とは異なる遺伝子のセットが発現して組織・細胞の癌化に寄与していると考えられている。したがって、正常組織とは異なる発現をする遺伝子は癌関連遺伝子である。正常な組織と比較して癌組織で発現変化する遺伝子を探索した。

乳がん由来のライブラリー (TBAES) と、正常な乳房由来のライブラリー (BEAST) のcDNAを解析して比較した結果 (表 9)、両者で発現変化のある遺伝子は以下の 4 クローンであった。

TBAES2004105, TBAES2007379, TBAES2007481, TBAES2008133

【 0 1 2 9 】

肝臓がん由来のライブラリー (TLIVE) と、正常な肝臓由来のライブラリー (LIVER) のcDNAを解析して比較した結果 (表 10)、両者で発現変化のある遺伝子は以下の 2 クローンであった。

THYMU3046350, TLIVE2007192

【 0 1 3 0 】

子宮がん由来のライブラリー (TUTER) と、正常な子宮由来のライブラリー (UTERU) のcDNAを解析して比較した結果 (表 1 1)、両者で発現変化のある遺伝子は以下の 3 クローンであった。

UTERU2017492, UTERU2025415, UTERU2036507

【 0 1 3 1 】

舌がん由来のライブラリー (CTONG) と、正常な舌由来のライブラリー (NTONG) のcDNAを解析して比較した結果 (表 1 2)、両者で発現変化のある遺伝子は以下の 2 クローンであった。

CTONG2002832, CTONG2003764

これらの遺伝子は、癌に関する遺伝子である。

【 0 1 3 2 】

また、発生や分化に関連する遺伝子を調べる方法として、発生・分化途中の組織・細胞と、成体の組織細胞において遺伝子発現量の違いを調べる発現頻度解析がある。組織の発生・分化に関する遺伝子は、その組織の構築と機能発現に関する遺伝子であり、傷害のある組織を任意に再生せしめる再生医学に利用可能な有用な遺伝子である。

【 0 1 3 3 】

先に記した1,402,070個のクローンの塩基配列のデータベースを基にして得た遺伝子発現頻度情報を用いて、発生・分化途中の組織・細胞と成体の組織・細胞とを比較して遺伝子発現頻度に変化のある遺伝子を解析した。

【 0 1 3 4 】

胎児の脳由来のライブラリー (FCBBF, FEBRAまたはOCBBF) と成体の脳由来のライブラリー (BRACE, BRALZ, BRAMY, BRAWH, BRCAN, BRCOC, BRHIP, BRSSN, BRSTNまたはBRTHA) のcDNAを解析し、胎児と成体で比較した結果 (表 1 3)、両者で発現変化のある遺伝子は以下の 1 3 クローンであった。

BRACE3002184, BRACE3026993, BRACE3046450, BRAMY3008096, BRAWH3013197, BRAWH3028645, BRTHA3004432, BRTHA3024233, CTONG2002832, TESTI4002988, THYM U3046350, TRACH1000193, TRACH3019290

これらの遺伝子は組織・細胞の再生に関する遺伝子である。

【0 1 3 5】

本発明のcDNAがコードしている蛋白質が、例えば、細胞の増殖・分化などの細胞状態を制御する因子である場合には、以下のようにして医薬品開発を行うことができる。ある種の細胞に、本発明によって提供される蛋白質や抗体を細胞内にマイクロインジェクションすることによって、細胞の増殖・分化などの細胞状態変化や、細胞内の特定の遺伝子の活性化または抑制を指標に低分子化合物等をスクリーニングすることができる。このスクリーニングは、例えば、以下のように行うことができる。

【0 1 3 6】

まず、本発明の蛋白質を発現させ組換え蛋白質の精製品を取得する。次いで、その精製蛋白質を、各種細胞株または初代培養細胞の細胞内にマイクロインジェクションして、増殖・分化などの細胞の変化を調べる。または、ある特定の細胞状態変化に作用することが知られている遺伝子の誘導をmRNA量、蛋白質量で検出する。あるいは、ある特定の細胞状態変化に影響を与えることが知られている遺伝子産物（蛋白質）の働きにより変化した細胞内の物質（低分子化合物など）量で検出する。そのときに培養液等に活性をスクリーニングしたい物質（低分子でも高分子でも可能）を添加しておくことにより、細胞状態の変化に及ぼす影響を指標にスクリーニングできる。

【0 1 3 7】

マイクロインジェクションしなくとも、本発明で取得した遺伝子を導入した形質転換細胞株を用いてのスクリーニングが可能である。本発明で取得した遺伝子産物が特定の細胞状態変化に作用することが明らかになった場合には、該遺伝子産物の変化を指標にスクリーニングできる。このようなスクリーニングにより、本発明による蛋白質が細胞状態、機能を制御するのを活性化または抑制する物質が開発されれば、医薬品への応用が考えられる。

【0 1 3 8】

また、本発明のcDNAがコードしている蛋白質が、例えば、分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質、疾患関連

蛋白質については、それぞれの蛋白質を用いた機能の解析に基づいて、例えば以下のようにして医薬品開発を行うことができる。

【0 1 3 9】

膜蛋白質の場合、細胞上に発現して受容体やリガンドとして機能する蛋白質である可能性が高い。したがって、本発明によって提供される膜蛋白質を、公知の、あるいは新規なリガンドや受容体との結合活性に基づいてスクリーニングすれば、新たなリガンドー受容体の関係を見出すことができる。スクリーニングは公知の方法に従って行うことができる。

たとえば次のようにして本発明の蛋白質に対するリガンドをスクリーニングすることができる。すなわち（a）本発明の蛋白質またはその部分ペプチド、またはそれらを発現する細胞に被検試料を接触させる工程、および（b）該蛋白質、該ペプチドまたは該細胞に結合する被検試料を選択する工程、とによって特定の蛋白質に結合するリガンドのスクリーニングが可能となる。

【0 1 4 0】

一方、例えば、以下のようにして本発明の蛋白質の受容体を発現する細胞をスクリーニングすることもできる。すなわち、（a）本発明の蛋白質またはその部分ペプチドに被検細胞試料を接触させる工程、および（b）該蛋白質またはその部分ペプチドに結合する細胞を選択する工程、とによって特定の蛋白質に結合する受容体のスクリーニングが可能である。

このスクリーニングは、例えば、以下のように行うことが可能である。まず、本発明の蛋白質を発現させ組換え蛋白質の精製品を取得する。次いで、その精製蛋白質を標識し、各種細胞株または初代培養細胞に対して結合アッセイを行い、これにより受容体を発現している細胞を選定する（本庶・新井・谷口・村松編 新生化学実験講座 7 増殖分化因子とその受容体 p203-236（1991）東京化学同人）。標識としては、 ^{125}I などのRI標識のほか、酵素（アルカリホスファターゼ等）標識も可能である。

また、本発明の蛋白質を標識せずに用いて、本発明の蛋白質と受容体を発現している細胞とを結合させた後に、本発明の蛋白質に対する抗体を標識して用いて検出することも考えられる。上記スクリーニングにより得られた本発明の蛋白質

の受容体を発現する細胞は、後述するように該受容体のアゴニストやアンタゴニストのスクリーニングに用いることが可能である。

【0141】

上記のスクリーニングにより、本発明の蛋白質が結合するリガンドや、本発明の蛋白質の受容体やその受容体を発現する細胞が得られれば、それらリガンドや受容体と結合する化合物のスクリーニングが可能となる。またそれらの結合活性を指標に、両者の結合を阻害する化合物（例えば、受容体アゴニストやアンタゴニスト）のスクリーニングが可能となる。

このスクリーニング方法は、本発明の蛋白質が受容体である場合は、（a）被検試料の存在下で、本発明の蛋白質または本発明の蛋白質を発現する細胞とリガンドとを接触させる工程、（b）該蛋白質または該蛋白質を発現する細胞とリガンドとの結合活性を検出する工程、および（c）被検試料非存在下において検出した場合と比較して該結合活性を低下させる化合物を選択する工程、とを含む。また本発明の蛋白質がリガンドである場合には、（a）被検試料の存在下で、本発明の蛋白質を該蛋白質の受容体または該受容体を発現する細胞に接触させる工程、（b）該蛋白質とその受容体または該受容体を発現する細胞との結合活性を検出する工程、および（c）被検試料非存在下において検出した場合と比較して、該結合活性を低下させる化合物を選択する工程、を含む。

スクリーニングに用いる被検試料としては、例えば、細胞抽出液、遺伝子ライブラリーの発現産物、合成低分子化合物、合成ペプチド、天然化合物などが挙げられるが、これらに制限されない。また、本発明の蛋白質との結合活性を指標とした上記のスクリーニングにより単離された化合物を被検試料として用いることも可能である。

【0142】

このスクリーニングにより単離される化合物は、本発明の蛋白質自体または本発明の蛋白質に対する受容体のアゴニストやアンタゴニストの候補となる。本発明の受容体とリガンドとの結合活性の低下によるリン酸化などの細胞内シグナルの変化をもとに、得られた化合物が本発明の蛋白質の受容体のアゴニストであるかアンタゴニストであるかを判定することができる。また、スクリーニングによ

って得られる化合物は、生体内において、本発明の蛋白質と相互作用する分子(受容体も含む)との該相互作用を阻害する化合物の候補ともなる。本発明の蛋白質、または本発明の蛋白質に結合する受容体、またはリガンド、更にはこれらの化合物は、本発明の蛋白質が関連する疾患の予防薬や治療薬への応用、または本発明の蛋白質が関連する疾患の検査薬への応用などが考えられる。

【0 1 4 3】

分泌蛋白質の場合、細胞の増殖・分化などの細胞状態を制御する因子の可能性がある。新たな細胞状態を制御する因子は、ある種の細胞に、本発明によって提供される分泌蛋白質を加えることによって、細胞の増殖・分化などの細胞の状態変化や、細胞内の特定の遺伝子の活性化を指標にスクリーニングすることにより見出すことができる。

このスクリーニングは、例えば、以下のように行うことが可能である。まず、本発明の蛋白質を発現させ組換え蛋白質の精製品を取得する。次いで、その精製蛋白質を、各種細胞株または初代培養細胞に添加して、増殖・分化などの細胞の変化を調べる。または、ある特定の細胞状態の変化に影響を与えることが知られている遺伝子の誘導をmRNA量、蛋白質量で検出する。あるいはある特定の細胞状態変化に影響を与えることが知られている遺伝子産物(蛋白質)の働きにより変化した細胞内の物質(低分子化合物など)量で検出する。

【0 1 4 4】

このようなスクリーニングにより、本発明による蛋白質が細胞状態、機能を制御するとなれば、本発明の蛋白質は、関連した疾患に対して、そのまま、あるいは一部適した状態に改変して、医薬品や検査薬への応用が考えられる。

また、先に膜タンパクについて記述したように、本発明によって提供される分泌蛋白質を用いて、公知の、あるいは新規なリガンドや受容体との結合活性に基づいてスクリーニングすれば、新たなリガンド-受容体の関係を見出すことができ、同様の方法でアゴニスト、アンタゴニストの判定が可能となる。こうして得られる化合物は、生体内において、本発明の蛋白質と相互作用する分子(受容体も含む)との該相互作用を阻害する化合物の候補ともなる。これら化合物は、本発明の蛋白質が関連する疾患の予防薬や治療薬、あるいは検査薬への応用が考え

られる。

【0145】

シグナル伝達関連蛋白質、転写関連蛋白質の場合は、細胞内外の刺激に反応して、ある蛋白質や遺伝子に作用する因子の可能性がある。新たな蛋白質、遺伝子に作用する因子は、ある種の細胞に、本発明によって提供される蛋白質を発現させることによって、細胞内の特定の遺伝子や蛋白質の活性化を指標にスクリーニングすることにより見出すことができる。

このスクリーニングは、例えば、以下のように行うことが可能である。まず、本発明の蛋白質を発現した形質転換細胞株を取得する。次いで、その形質転換細胞株と、もとの未形質転換細胞株とにおいて、ある特定の遺伝子の変化をmRNA量、蛋白質量で検出する。あるいは、ある特定の遺伝子産物（蛋白質）の働きにより変化した細胞内の物質（低分子化合物など）量で検出する。さらには、ある特定の遺伝子の発現調節領域とマーカー遺伝子（ルシフェラーゼ、 β -ガラクトシダーゼ等）の融合遺伝子を導入した細胞に、本発明によって提供される蛋白質を同時に発現させることによって、特定の遺伝子の発現の変化を、マーカー遺伝子産物（蛋白質）由来の活性で判定する。

【0146】

このようなスクリーニングにより、影響を受けた蛋白質や遺伝子が疾患に関連していた場合、本発明による蛋白質を利用し、直接的に、または、間接的に、その発現や活性調節を行う化合物や遺伝子のスクリーニングが可能となる。

例えば、まず、本発明の蛋白質を発現させ組換え蛋白質の精製品を取得する。次に影響を受けた蛋白質や遺伝子を精製し、その結合を調べる。または、予め阻害剤の候補となる化合物を加えておいた後、それら結合の変化を調べる。あるいは、例えば他遺伝子の発現調節を行う本発明の蛋白質をコードする遺伝子の5'上流転写調節領域を取得し、マーカー遺伝子と融合した遺伝子を導入した細胞に、化合物などを添加して、当該遺伝子の発現を制御する因子を見いだす。

【0147】

このようなスクリーニングによって得られた化合物は、本発明による蛋白質が関連した疾患に対して医薬品への応用が考えられる。スクリーニングによって得

られた制御因子が蛋白質であっても、同様に、その蛋白質の発現・活性に本来ない影響を与える化合物があれば、その化合物は、本発明による蛋白質が関連した疾患に対して医薬品への応用が考えられる。

分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質、疾患関連蛋白質のいずれの場合でも、本発明による蛋白質が酵素としての活性を有するとなれば、本発明によって提供される蛋白質に化合物を適当な条件下で添加し、化合物の変化を指標にスクリーニングすれば可能である。また、この活性を指標に本発明による蛋白質の活性を阻害する化合物のスクリーニングも可能である。

【0148】

このスクリーニングは、例えば、以下のように行うことが可能である。まず、本発明の蛋白質を発現させ組換え蛋白質の精製品を取得する。次いで、その精製蛋白質に、化合物を添加して、化合物量および反応生成物量を調べる。または、予め阻害剤の候補となる化合物を加えておいた後、精製蛋白質と反応する化合物（基質）を加えて、その基質量および反応生成物量の変化を調べる。

このようなスクリーニングにより、得られた化合物は、本発明の蛋白質が関連した疾患に対して、医薬品への応用が考えられる。また本発明の蛋白質が生体において正常に機能しているかどうかを調べるなどの検査への応用が考えられる。

【0149】

本発明の分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質が、新たな疾患関連蛋白質であるかは、上記に挙げた以外に、本発明による蛋白質を発現して得られた特異認識抗体を用いて、特定の疾患と蛋白質の発現量や活性との相関を知ることができる。あるいは、「Method in Molecular Biology」(Humana Press社)シリーズの『Molecular Diagnosis of Genetic Diseases』(Rob Elles編、1996)を参考に解析が可能である。

疾患関連蛋白質は、前述のようなスクリーニングの対象となり、その発現・活性を制御する薬剤の開発に有用である。また、関連した疾患の診断マーカー、あるいは遺伝子治療のターゲットになるなど医療産業上、有用である。

【0150】

以上により単離された化合物を医薬品として用いる場合には、単離された化合物自体を直接患者に投与する以外に、公知の製剤学的方法により製剤化して投与を行うことも可能である。例えば、薬理学上許容される担体もしくは媒体、具体的には、滅菌水や生理食塩水、植物油、乳化剤、懸濁剤などと適宜組み合わせ、製剤化して投与することが考えられる。患者への投与は、例えば、動脈内注射、静脈内注射、皮下注射など当業者に公知の方法により行いうる。投与量は、患者の体重や年齢、投与方法などにより変動するが、当業者であれば適当な投与量を適宜選択することが可能である。また、該化合物がDNAによりコードされるものであれば、該DNAを遺伝子治療用ベクターに組み込み、遺伝子治療を行うことも考えられる。投与量、投与方法は、患者の体重や年齢、症状などにより変動するが、当業者であれば適宜選択することが可能である。

【0 1 5 1】

更に本発明は、表 1 に記載された塩基配列および／またはアミノ酸配列から選択された少なくとも 1 つを含む全長cDNAおよび／または蛋白質データベースに関する。データベースとは、塩基配列情報を検索可能な機械可読式の情報として蓄積した情報の集合を意味する。本発明のデータベースは、本発明によって提供されるcDNAの塩基配列の少なくとも一つを含む。本発明のデータベースは、本発明によって提供されるcDNAのみから構成されていても良いし、公知の全長cDNAやEST等の塩基配列情報をも含むものであることができる。本発明のデータベースには、塩基配列情報のみならず、本発明によって明らかにされた遺伝子の機能情報や、その全長cDNAを保持したクローンの名称などの付随した情報を合わせて記録したり、あるいはリンクさせておくことができる。

【0 1 5 2】

本発明のデータベースは、遺伝子断片の情報に基づく、遺伝子全長の取得に有用である。本発明に基づくデータベースは、いずれも全長cDNAの塩基配列情報からなっている。したがって、ディファレンシャルディスプレイ法や、サブトラクション法によって得られた遺伝子断片の塩基配列を、このデータベースの情報に照合すれば、断片の塩基配列に基づいて遺伝子の全長塩基配列を明らかにすることができる。

しかも本発明のデータベースを構成する全長cDNAの配列情報は、全長であることのみならず、遺伝子の発現頻度情報や、公知の遺伝子や蛋白質との相同性検索結果を伴っていることから、遺伝子断片の機能解析を飛躍的に迅速化する。更に本発明のデータベースは、ヒトの遺伝子に関する情報を蓄積しているものであることから、他の種から単離された遺伝子の塩基配列情報に基づくヒトのホモログの単離に有用である。

【0153】

現在では、ディファレンシャルディスプレイ法や、サブトラクション法によって、さまざまな遺伝子断片情報を得ることができる。一般にこれらの遺伝子断片は、その全長を取得するためのツールとして用いられる。遺伝子断片が公知の遺伝子のものであれば、公知のデータベースとの照合によって、その全長を明らかにすることは容易である。しかし、公知の遺伝子データベースに一致する塩基配列を見出せない場合には、全長cDNAのクローニングを行わなければならない。これらの断片情報に基づいて全長塩基配列を取得する工程は、しばしば困難を伴う。遺伝子の全長を取得しない限り、その遺伝子がコードする蛋白質のアミノ酸配列は明らかにできない。したがって、本発明のデータベースは、公知の遺伝子のデータベースでは解明することのできない、遺伝子断片に対応する全長cDNAの特定に貢献する。

【0154】

【実施例】

実施例 1. オリゴキャップ法によるcDNAライブラリーの作製

(1) mRNA抽出と購入

ヒト組織（下記に示す）より、文献（J. Sambrook, E. F. Fritsch & T. Maniatis, Molecular Cloning Second edition, Cold Spring harbor Laboratory Press, 1989）記載の方法により全RNAとしてmRNAを抽出した。また、ヒト培養細胞やヒト初代培養細胞（下記に示す）をカタログ記載の方法で培養後、文献（J. Sambrook, E. F. Fritsch & T. Maniatis, Molecular Cloning Second edition, Cold Spring harbor Laboratory Press, 1989）記載の方法により全RNAとしてmRNAを抽出した。

【 0 1 5 5 】

以下にライブラリー名とその由来の関係を、「ライブラリー名：由来」の順に示した。サブトラクションしたものについては、サブトラクトライブラリーの作り方も示した。

< ヒト組織より mRNA 抽出 >

NTONG：正常舌(Tongue)；

CTONG：舌癌(Tongue, Cancer)；

FCBBF：胎児脳(Brain, Fetal)；

OCBBF：胎児脳(Brain, Fetal)；

PLACE：胎盤(Placenta)；

SYNOV：滑膜組織(Synovial membrane tissue from rheumatoid arthritis)；

CORDB：臍帯血(Cord blood)。

【 0 1 5 6 】

< 培養細胞より mRNA 抽出 >

BNGH4：H4細胞(ATCC #HTB-148)；

IMR32：IMR32細胞(ATCC #CCL-127)；

SKNMC：SK-N-MC細胞(ATCC #HTB-10)；

3NB69：NB69細胞(RCB #RCB0480)；

BGGI1：GI1細胞(RCB #RCB0763)；

NB9N4：NB9細胞(RCB #RCB0477)；

SKNSH：SK-N-SH細胞(RCB #RCB0426)；

AHMSC：HMSC細胞((間葉細胞, Human mesenchymal cell)；

CHONS：軟骨細胞(Chondrocyte)；

ERLTF：TF-1細胞((赤白血病細胞, erythroleukemia)；

HELAC：HeLa細胞；

JCMLC：白血病細胞(Leukemia, myelogenous)；

MESTC：間葉系幹細胞((Mesenchyme stem cell)；

NIESE：間葉系幹細胞(Mesenchymal stem cell)；

NCRRM：胎生期癌細胞(Embryonal carcinoma)；

NCRRP：胎生期癌細胞(Embryonal carcinoma)をレチノイン酸(RA)処理誘導；
T1ESE：間葉系幹細胞(Mesenchymal stem cell) をトリコスタチンと 5 アザシチジン処理誘導；
NT2RM：NT2細胞(STARATAGENE #204101)；
NT2RP：NT2細胞をレチノイン酸(RA)処理誘導5週間；
NT2RI：NT2細胞をRA処理誘導5週間後、生育阻害剤処理2週間；
NT2NE：NT2細胞をRA処理と生育阻害剤処理により神経分化後、神経を濃縮回収(N T2 Neuron)；
NTISM：NT2細胞(STARATAGENE #204101)をRA処理誘導5週間後、生育阻害剤処理を2週間したmRNAから作製したcDNAライブラリーから、未分化NT2細胞のmRNAと重複する c DNAをSubtract Kit (Invitrogen #K4320-01)を用いてサブトラクトしたライブラリー (NT2RI-NT2RM)。

RCBは、理化学研究所ジーンバンク・細胞開発銀行より分譲をうけたものであり、ATCCは、American Type Culture Collectionより分譲をうけたものである。

【 0 1 5 7 】

<初代培養細胞よりmRNA抽出>

ASTRO：正常神経膠星状細胞(Normal Human Astrocyte) NHA5732, 宝酒造 #CC2565；

DFNES：新生児正常皮膚繊維芽細胞(Normal Human Dermal Fibroblasts (Neonatal Skin)；NHDF-Neo) NHDF2564, 宝酒造 #CC2509；

MESAN：正常メサンギウム細胞(Normal human mesangial cells) NHMC56046-2, 宝酒造 #CC2559；

NHNPC：正常神経前駆細胞(Normal human neural progenitor cells) NHNP5958, 宝酒造 #CC2599；

PEBLM：正常末梢血単核細胞(Human peripheral blood mononuclear cells) HPBM C5939, 宝酒造 #CC2702；

HSYRA：滑膜細胞HS-RA(Human synoviocytes from rheumatoid arthritis), 東洋紡 #T404K-05；

PUAEN：正常肺動脈内皮細胞(Human pulmonary artery endothelial cells), 東

洋紡 #T302K-05 ;

UMVEN : 正常臍帯静脈内皮細胞 (Human umbilical vein endothelial cells) HUVEC, 東洋紡 #T200K-05 ;

HCASM : 正常冠動脈平滑筋細胞 HCASMC (Human coronary artery smooth muscle cells), 東洋紡 #T305K-05 ;

HCHON : 正常軟骨細胞 HC (Human Chondrocytes), 東洋紡 #T402K-05 ;

HHDPC : 正常頭髮毛乳頭細胞 HDPC (Human dermal papilla cells), 東洋紡 #THPCK-001 ;

CD34C : CD34+細胞 (AllCells, LLC #CB14435M) ;

D30ST : CD34+細胞を破骨細胞分化因子 (ODF) 処理誘導3日間 ;

D60ST : CD34+細胞をODF処理誘導6日間 ;

D90ST : CD34+細胞をODF処理誘導9日間 ;

ACTVT : 活性化T細胞 (Activated T-cell) ;

LYMPB : リンパ芽球 (Lymphoblast, EB virus transferred B cell) ;

NETRP : 好中球 (Neutrophil)。

【 0 1 5 8 】

次いで、以下に示すヒト組織より全RNAとして抽出されたmRNAを購入した。以下にライブラリー名とその由来の関係を、「ライブラリー名：由来」の順に示した。サブトラクションしたものについては、サブトラクトライブラリーの作り方も示した。

<ヒト組織よりのmRNAを全RNAで購入>

ADRGL : 副腎 (Adrenal gland), CLONTECH #64016-1 ;

BRACE : 小脳 (Brain, cerebellum), CLONTECH #64035-1 ;

BRAWH : 全脳 (Brain, whole), CLONTECH #64020-1 ;

FEBRA : 胎児脳 (Brain, Fetal), CLONTECH #64019-1 ;

FELIV : 胎児肝臓 (Liver, Fetal), CLONTECH #64018-1 ;

HEART : 心臓 (Heart), CLONTECH #64025-1 ;

HLUNG : 肺 (Lung), CLONTECH #64023-1 ;

KIDNE : 腎臓 (Kidney), CLONTECH #64030-1 ;

LIVER：肝臓(Liver), CLONTECH #64022-1；
MAMGL：乳腺(Mammary Gland), CLONTECH #64037-1；
PANCR：膵臓(Pancreas), CLONTECH #64031-1；
PROST：前立腺(Prostate), CLONTECH #64038-1；
SALGL：唾液腺(Salivary Gland), CLONTECH #64026-1；
SKMUS：骨格筋(Skeletal Muscle), CLONTECH #64033-1；
SMINT：小腸(Small Intestine), CLONTECH #64039-1；
SPLEN：脾臓(Spleen), CLONTECH #64034-1；
STOMA：胃(Stomach), CLONTECH #64090-1；
TBAES：乳癌(Breast, Tumor), CLONTECH #64015-1；
TCERX：子宮頸管癌(Cervix, Tumor), CLONTECH #64010-1；
TCOLN：結腸癌(Colon, Tumor), CLONTECH #64014-1；
TESTI：精巣(Testis), CLONTECH #64027-1；
THYMU：胸腺(Thymus), CLONTECH #64028-1；
TLUNG：肺癌(Lung, Tumor), CLONTECH #64013-1；
TOVAR：卵巢癌(Ovary, Tumor), CLONTECH #64011-1；
TRACH：気管(Trachea), CLONTECH #64091-1；
TUTER：子宮癌(Uterus, Tumor), CLONTECH #64008-1；
UTERU：子宮(Uterus), CLONTECH #64029-1；
ADIPS：脂肪組織(Adipose), Invitrogen #D6005-01；
BLADE：膀胱(Bladder), Invitrogen #D6020-01；
BRALZ：アルツハイマー患者大脳皮質(Brain, cortex, Alzheimer), Invitrogen #D6830-01；
CERVX：子宮頸管(Cervix), Invitrogen #D6047-01；
COLON：結腸(Colon), Invitrogen #D6050-0；
NESOP：食道(Esophagus), Invitrogen #D6060-01；
PERIC：心膜(Pericardium), Invitrogen #D6105-01；
RECTM：直腸(Rectum), Invitrogen #D6110-01；
TESOP：食道癌(Esophageal, Tumor), Invitrogen #D6860-01；

TKIDN：腎臓癌(Kidney, Tumor), Invitrogen #D6870-01；
TLIVE：肝臓癌(Liver, Tumor), Invitrogen #D6880-01；
TSTOM：胃癌(Stomach, Tumor), Invitrogen #D6920-01；
BEAST：成人乳房(Adult Breast), STARATAGENE #735044；
FEHRT：胎児心臓(Heart, Fetal), STARATAGENE #738012；
FEKID：胎児腎臓(Kidney, Fetal), STARATAGENE #738014；
FELNG：胎児肺(Lung, Fetal), STARATAGENE #738020；
NOVAR：成人卵巣(Adult Ovary), STARATAGENE #735260；
BRASW：アルツハイマー患者大脳皮質組織 [BRALZ：アルツハイマー患者大脳皮質
(Brain, cortex, Alzheimer), Invitrogen #D6830-01] のmRNAから作製したcDNA
ライブラリーから、全脳組織 [BRAWH：全脳(Brain, whole), CLONTECH #64020-1
] のmRNAと重複する cDNAをSubtract Kit (Invitrogen #K4320-01)を用いてサブ
トラクトしたライブラリー (BRALZ-BRAWH)。

【 0 1 5 9 】

さらに、次に示すヒト組織よりポリA(+) RNAとして抽出・精製されたmRNAを購入した。各組織由来のポリA(+) RNAに、ポリA(-)RNAを混ぜたRNAからcDNAライブラリーを作製した。ポリA(-)RNAは、全脳(Brain, whole), CLONTECH #64020-1の全RNAからポリA(+)RNAをオリゴdTセルロースで除くことにより調製した。以下にライブラリー名とその由来の関係を、「ライブラリー名：由来」の順に示した。

<ヒト組織よりのmRNAをポリA(+) RNAで購入>

BRAMY：扁桃(Brain, amygdala), CLONTECH #6574-1；
BRCAN：尾状核(Brain, caudate nucleus), CLONTECH #6575-1；
BRCOC：脳梁(Brain, corpus callosum), CLONTECH #6577-1；
BRHIP：海馬(Brain, hippocampus), CLONTECH #6578-1；
BRSSN：黒質(Brain, substantia nigra), CLONTECH #6580-1；
BRSTN：視床下核(Brain, subthalamic nucleus), CLONTECH #6581-1；
BRTHA：視床(Brain, thalamus), CLONTECH #6582-1。

【 0 1 6 0 】

(2) cDNAライブラリーの作製

それぞれのRNAよりオリゴキャップ法[M. Maruyama and S. Sugano, Gene, 138: 171-174 (1994)]を改良した方法(WO 01/04286)によりcDNAライブラリーを作製した。Oligo-cap linker (agcaucgagu cgcccuuguu ggccuacugg/配列番号：79) およびOligo dT primer (gcggctgaag acggcctatg tggccttttt tttttttttt tt/配列番号：80)を用いて、WO 01/04286に記載したようにBAP (Bacterial Alkaline Phosphatase) 処理、TAP (Tobacco Acid Pyrophosphatase) 処理、RNAライゲーション、第一鎖cDNAの合成とRNAの除去を行った。次いで、5' (agcatcgagt cgcccttggt g/配列番号：81) と3' (gcggctgaag acggcctatg t/配列番号：82) のPCRプライマーを用いPCR (polymerase chain reaction)により2本鎖cDNAに変換し、SfiIで切断した。次いで、通常は2kb以上（場合によっては3kb以上）に分画したcDNA断片をDraIIIで切断したベクターpME18SFL3 (図1) (GenBank AB009864, Expression vector) にcDNAの方向性を決めてクローニングし、cDNAライブラリーを作製した。

全長cDNA配列解析に用いたcDNAライブラリー名とその由来の関係を表2に示した。

【0161】

【表2】

ライブラリー名	タイプ	由来など
ADRGL	組織	副腎(Adrenal gland) (CLONTECH #64016-1)
BRACE	組織	小脳(Brain, cerebellum) (CLONTECH #64035-1)
BRAMY	組織	扁桃(Brain, amygdala) (CLONTECH #6574-1)
BRAWH	組織	全脳(Brain, whole) (CLONTECH #64020-1)
BRCAN	組織	尾状核(Brain, caudate nucleus) (CLONTECH #6575-1)
BRHIP	組織	海馬(Brain, hippocampus) (CLONTECH #6578-1)
BRTHA	組織	視床(Brain, thalamus) (CLONTECH #6582-1)
CTONG	組織	舌癌(Tongue, Cancer)
PLACE	組織	胎盤(Placenta)
PROST	組織	前立腺(Prostate) (CLONTECH #64038-1)
TBAES	組織	乳癌(Breast, Tumor) (CLONTECH #64015-1)
TESTI	組織	精巣(Testis) (CLONTECH #64027-1)
THYMU	組織	胸腺(Thymus) (CLONTECH #64028-1)
TLIVE	組織	肝臓癌(Liver, Tumor) (Invitrogen #D6880-01)
TRACH	組織	気管(Trachea) (CLONTECH #64091-1)
UTERU	組織	子宮(Uterus) (CLONTECH #64029-1)

【0162】

オリゴキャップ法を改良した方法で作製した高全長率cDNAライブラリー（既知

mRNAのタンパク質コード領域を指標にして算出した各cDNAライブラリーの5'端の全長率は平均90%)は、真核細胞での発現が可能な発現ベクターpME18SFL3を用いて作製した。pME18SFL3にはクローニング部位の上流にSR α プロモーターとSV40 small tイントロンが組み込まれており、またその下流にはSV40ポリA付加シグナル配列が挿入されている。pME18SFL3のクローン化部位は非対称性のDraII Iサイトとなっており、cDNA断片の末端にはこれと相補的なSfiI部位を付加しているため、クローン化したcDNA断片はSR α プロモーターの下流に一方向性に挿入される。したがって、全長cDNAを含むクローンでは、得られたプラスミドをそのままCOS細胞などに導入することにより、一過的に遺伝子を発現させることが可能である。すなわち、非常に容易に、遺伝子産物である蛋白質として、あるいはそれらの生物学的活性として実験的に解析することが可能となっている。

【0163】

(3) オリゴキャップ法で作製したcDNAライブラリーからのクローンの5'-末端の全長性の評価

これらより得たクローンのプラスミドDNAについて、cDNAの5'端(一部のクローンについては3'端も)の塩基配列をDNAシーケンシング試薬(BigDye Terminator Cycle Sequencing FS Ready Reaction Kit, PE Biosystems社製)を用い、マニュアルに従ってシーケンシング反応後、DNAシーケンサー(ABI PRISM 3700, PE Biosystems社製)でDNA塩基配列を解析した。得られたデータをデータベース化した。

【0164】

オリゴキャップ法を改良した方法で作製したヒトcDNAライブラリーの約111万クローンの5'-末端の全長率を次の方法で求めた。公共データベース中のヒト既知mRNAと5'-末端配列が一致する全クローンについて、公共データベース中の既知mRNA配列より長く5'-末端が伸びている場合、または5'-末端は短いが翻訳開始コドンは有している場合を「全長」と判断し、翻訳開始コドンを含んでいない場合を「非全長」と判断した。これをもとに5'-末端の全長率[全長クローン数/(全長クローン数+非全長クローン数)]を計算した。この結果、5'-末端の全長率は、90%であった。この結果より、オリゴキャップ法で取得したヒトcDNAラ

イブラリーからのクローンの5'-端配列の全長率が非常に高いことが分かった。

【0 1 6 5】

実施例 2. cDNAクローン末端配列解析と全長塩基配列解析クローンの選択

各cDNAライブラリーより得たクローンのプラスミドDNAについて、cDNAの5'末端の塩基配列をDNAシーケンシング試薬 (Dye Terminator Cycle Sequencing FS Ready Reaction Kit, dRhodamine Terminator Cycle Sequencing FS Ready Reaction KitまたはBigDye Terminator Cycle Sequencing FS Ready Reaction Kit, PE Biosystems社製) を用い、マニュアルに従ってシーケンシング反応後、DNAシーケンサー (ABI PRISM 3700, PE Biosystems社製) で解析した。得られたデータについてはデータベース化を行った。

【0 1 6 6】

解析されたcDNAクローンの5'末端配列については、GenBank、UniGeneのcomplete cdsの表記があるデータを対象にしたBLASTによる相同性検索を行い、ヒトのmRNA配列に同一なものは除いた。次にクラスタリングを行い、相同性90%以上かつコンセンサス配列が50塩基対以上の場合、同一グループと見なし、グループを形成させた。グループ内の、より5'-側に長いクローンを選択し、選択されたクローンについては必要に応じ3'末端配列を5'末端配列と同様の方法で解析取得した。取得された末端配列のデータを解析し、5'末端と3'末端の配列でコンティグを作るクローンは除いた。更に再度前記と同様にBLASTによる相同性検索によりヒトのmRNA配列 (特許化または特許出願された配列を含む) に同一なものは除いた。こうして選択したクローンより全長塩基配列解析を行うクローンを得た。

【0 1 6 7】

実施例 3. 全長塩基配列解析

全長塩基配列解析に選抜されたクローンについて各々全長cDNAの塩基配列を決定した。塩基配列は、主にカスタム合成DNAプライマーを用いたダイデオキシターミネーター法によるプライマーウォーキング法によって決定した。すなわち、カスタム合成DNAプライマーを用い、PE Biosystem社製のDNAシーケンシング試薬でマニュアルに従ってシーケンシング反応後、同社製のシーケンサーを用いてDNA塩基配列を解析した。一部のクローンについては、Licor社製DNAシーケンサー

も利用した。

また、一部のクローンについてはカスタムプライマーを用いずcDNA が含まれるプラスミドをランダムに切断するショットガン法を用いて同様にDNAシーケンサーでDNA塩基配列を決定した。全長塩基配列は上記方法により決定された部分塩基配列を完全にオーバーラップさせ最終的に確定した。

次に、決定された全長塩基配列から、蛋白質への翻訳領域を推定しアミノ酸配列を求めた。それぞれに対応する配列番号を表1に示す。

【0168】

実施例4. 相同性検索による機能予測

決定された塩基配列および蛋白質をコードすると推定されるORF部分のアミノ酸配列についてSwissProt、Ref-Seq、nrに対するBLAST検索を行った。P値またはE値が 10^{-4} 以下であり、かつアミノ酸データベースを対象にした解析においてはコンセンサス長×相同性 = 30以上のBLAST検索ヒットデータの中から、相同性がより高く、塩基配列及び推定アミノ酸配列に対して機能の予測が比較的容易なヒットデータの中から代表的なものを選択し、相同性検索結果データとして明細書の最後に示した。したがって示したデータはあくまで代表的なものであり、各クローンに相同性を示す分子が、これのみに限定されるというわけではない。また、一部のクローンにおいて、先に記した条件にあてはまらないBLAST検索ヒットデータについては示さなかった。

【0169】

実施例5. 推定アミノ酸配列に対するシグナル配列、膜貫通領域および機能ドメインの検索

全長塩基配列から推定されたアミノ酸配列に対して、アミノ末端のシグナル配列の有無と膜貫通領域の有無を予測、さらに蛋白質の機能ドメイン（モチーフ）検索を行った。アミノ末端のシグナル配列についてはPSORT [K. Nakai & M. Kanehisa, Genomics, 14: 897-911 (1992)]を、膜貫通領域についてはSOSUI [T. Hirokawa et.al. Bioinformatics, 14: 378-379 (1998)]（三井情報開発株式会社販売）を用いて解析を行った。機能ドメインの検索についてはPfam (Version 5.5) (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>) を用いた。PSORTや

SOSUIにより、アミノ末端のシグナル配列や膜貫通領域が予測されたアミノ酸配列は分泌、膜蛋白質であると予測された。また、Pfamによる機能ドメイン検索において、ある機能ドメインにヒットしたアミノ酸配列はヒットデータをもとに、例えばPROSITE(<http://www.expasy.ch/cgi-bin/prosite-list.pl>)にある機能カテゴリー分類を参照にしてその蛋白質の機能予測することができる。また、PROSITEでの機能ドメインの検索も可能である。

各ソフトウェアによる検索結果を以下に示す。

【 0 1 7 0 】

PSORTにより推定アミノ酸配列にシグナル配列を検出されたクローンは、以下の9クローンであった。

ADRGL2011190, PROST2010326, TBAES2004105, TBAES2007379, TBAES2008133, THYMU3044175, TLIVE2007192, UTERU2025415, UTERU2036507

【 0 1 7 1 】

SOSUIにより推定アミノ酸配列に膜貫通領域を検出されたクローンは、以下の14クローンであった。数字は推定アミノ酸配列中に検出された膜貫通領域の数を示している。検索結果は、クローン名と膜貫通領域の数を//で区切って示した。

BRAWH3013197//6, BRAWH3028645//7, BRAWH3046240//7, BRTHA3024233//3, PROST2010326//8, TBAES2004105//1, TBAES2007379//1, TBAES2007481//1, TESTI2043585//8, THYMU3024602//8, THYMU3044175//2, THYMU3046350//9, TLIVE2007192//2, TRACH3021066//2

【 0 1 7 2 】

Pfamにより推定アミノ酸配列に機能ドメインを検出されたクローンは以下の29クローンであった。検索結果は、クローン名//機能ドメイン名のように示し、複数の機能ドメインがヒットした場合には//で区切って並記した。なお同一の機能ドメインが複数ヒットした場合も省略せずに記載した。

ADRGL2011190// GAF domain// Transposase, Mutator family// 3'5'-cyclic nucleotide phosphodiesterase

BRACE3002184// KRAB box// Zinc finger, C2H2 type// Zinc finger, C2H2 typ

e// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2
type// Zinc finger, C2H2 type// Transcription factor S-II (TFIIS)// Zinc
finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Z
inc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type/
/ Zinc finger, C2H2 type// Zinc finger, C2H2 type

BRACE3026993// TSC-22/dip/bun family

BRACE3046450// PDZ domain (Also known as DHR or GLGF).// PDZ domain (Also known as DHR or GLGF).// PDZ domain (Also known as DHR or GLGF). 1// PDZ domain (Also known as DHR or GLGF).

BRAMY3008096// Eukaryotic protein kinase domain

BRAWH3013197// Receptor family ligand binding region// 7 transmembrane r
eceptor (metabotropic glutamate family)

BRAWH3028645// Latrophilin/CL-1-like GPS domain// FecCD transport family
// 7 transmembrane receptor (Secretin family)

BRAWH3046240// 7 transmembrane receptor (rhodopsin family)

BRCAN2019772// Zinc finger, C4 type (two domains)// Ligand-binding domain of nuclear hormone receptor

CTONG2002832// KRAB box// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc carboxypeptidase// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type// Zinc finger, C2H2 type

CTONG2003764// Phorbol esters/diacylglycerol binding domain (C1 domain)

PROST2010326// Latrophilin/CL-1-like GPS domain// Cytochrome C assembly
protein// 7 transmembrane receptor (Secretin family)

TBAES2004105// Thrombospondin type 1 domain

TBAES2007379// EGF-like domain// EGF-like domain// Trypsin Inhibitor like cysteine rich domain// EGF-like domain// EGF-like domain// Keratin, hi

【 0 1 7 3 】

全長塩基配列および推定アミノ酸配列のSwiss-Prot、nr、RefSeqの各データベースを対象に行った相同性検索の結果（相同性検索結果データ参照）から、クローン中にコードされる蛋白質の機能予測、カテゴリー分類を行った。

分泌・膜蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中に growth factor, cytokine, hormone, signal, transmembrane, membrane, extracellular matrix, receptor, G-protein coupled receptor, ionic channel, voltage-gated channel, calcium channel, cell adhesion, collagen, connective tissue 等、分泌・膜蛋白質と推定される記載があった、もしくはPsor

tとSOSUIによる推定ORFの解析の結果、シグナルシーケンスや膜貫通領域があったクローンである。

【0 1 7 5】

糖蛋白質関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中に glycoprotein 等、糖蛋白質関連蛋白質と推定される記載があったクローンである。

【0 1 7 6】

シグナル伝達関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中に serine/threonine-protein kinase, tyrosine-protein kinase, SH3 domain, SH2 domain等、シグナル伝達関連蛋白質と推定される記載があったクローンである。

【0 1 7 7】

転写関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中に transcription regulation, zinc finger, homeobox 等、転写関連蛋白質と推定される記載があったクローンである。

【0 1 7 8】

疾患関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中に disease mutation, syndrome 等、疾患関連蛋白質と推定される記載があった、あるいは全長塩基配列に対するSwiss-Protヒットデータ、及びGenBank、UniGeneヒットデータが、ヒトの遺伝子と疾患のデータベースであるOnline Mendelian Inheritance in Man (OMIM)に登録されている遺伝子、蛋白質であったクローンである。

【0 1 7 9】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にmetabolism, oxidoreductase, E.C.No. (Enzyme commission number)等、酵素・代謝関連蛋白質と推定される記載があったクローンである。

【0 1 8 0】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンとは、cell division, cell cycle, mitosis, chromosomal protein, cell growth, ap

optosis等、細胞分裂・増殖関連蛋白質と推定される記載があったクローンである。

【0181】

細胞骨格関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にstructural protein, cytoskeleton, actin-binding, microtubules等、細胞骨格関連蛋白質と推定される記載があったクローンである。

【0182】

核蛋白質・RNA合成関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にnuclear protein, RNA splicing, RNA processing, RNA helicase, polyadenylation等、核蛋白質・RNA合成関連蛋白質と推定される記載があったクローンである。

【0183】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にtranslation regulation, protein biosynthesis, amino-acid biosynthesis, ribosomal protein, protein transport, signal recognition particle等、蛋白質合成・輸送関連蛋白質と推定される記載があったクローンである。

【0184】

細胞防御関連蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にheat shock, DNA repair, DNA damage等、細胞防御関連蛋白質と推定される記載があったクローンである。

【0185】

発生・分化関連蛋白質のカテゴリーに属すると推定されたクローンとは、developmental protein等、発生・分化関連蛋白質と推定される記載があったクローンである。

【0186】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にDNA-binding, RNA-binding等と記載があったクローンである。

【0187】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンとは、ヒットデータ中にATP-binding, GTP-binding等と記載があったクローンである。

【0188】

この機能カテゴリー分類では一つのクローンが上記の複数のカテゴリーに該当する場合は、そのまま複数のカテゴリーに分類した。ただし、蛋白質の機能は必ずしも分類された機能カテゴリーに限定されるわけではなく、今後その他の機能も明らかになる可能性がある。

【0189】

分泌・膜蛋白質に属すると推定されたクローンは、以下の19クローンであった。

ADRL2011190, BRACE3046450, BRAWH3013197, BRAWH3028645, BRAWH3046240, BRTHA3024233, PROST2010326, TBAES2004105, TBAES2007379, TBAES2007481, TBAES2008133, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3044175, THYMU3046350, TLIVE2007192, UTERU2025415, UTERU2036507

【0190】

糖蛋白質関連蛋白質に属すると推定されたクローンは、以下の10クローンであった。

BRAWH3013197, BRAWH3046240, PROST2010326, TBAES2004105, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3046350, TRACH3021066, UTERU2036507

【0191】

シグナル伝達関連蛋白質に属すると推定されたクローンは、以下の4クローンであった。

BRHIP3030064, CTONG2003764, PLACE7013963, TRACH3021066

【0192】

転写関連蛋白質に属すると推定されたクローンは、以下の7クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【0193】

疾患関連蛋白質に属すると推定されたクローンは、以下の 2 3 クローンであった。

ADRGL2011190, BRACE3002184, BRACE3026993, BRACE3046450, BRAWH3013197, BRAWH3046240, BRCAN2019772, BRHIP3030064, BRTHA3024233, CTONG2002832, CTONG2003764, PROST2010326, TBAES2004105, TESTI2046536, TESTI4005158, TESTI4005500, THYMU3024602, THYMU3046350, TRACH1000193, TRACH3019290, TRACH3021066, UTERU2017492, UTERU2036507

【 0 1 9 4 】

また上記 2 3 クローンは全てが、Swiss-Prot ヒットデータ、及び nr、RefSeq ヒットデータが、ヒトの遺伝子と疾患のデータベースである Online Mendelian Inheritance in Man (OMIM) に登録されている遺伝子、蛋白質であった。以下にクローン名と OMIM Number を示す (クローン名の後ろのカッコ内は対象となった OMIM Number)。

ADRGL2011190 (602658), BRACE3002184 (603899), BRACE3026993 (601282), BRACE3046450 (603035), BRAWH3013197 (604100), BRAWH3046240 (602646), BRCAN2019772 (603849), BRHIP3030064 (605762), BRTHA3024233 (603088), CTONG2002832 (604668), CTONG2003764 (176977), PROST2010326 (604110), TBAES2004105 (605008), TESTI2046536 (182890; 182882), TESTI4005158 (601328), TESTI4005500 (602277), THYMU3024602 (600493), THYMU3046350 (603080), TRACH1000193 (139139), TRACH3019290 (603975), TRACH3021066 (164870), UTERU2017492 (602917), UTERU2036507 (605008)

【 0 1 9 5 】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 8 クローンであった。

ADRGL2011190, BRHIP3030064, CTONG2003764, PLACE7013963, TBAES2004105, THYMU3044175, TRACH3021066, UTERU2036507

【 0 1 9 6 】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

TBAES2007481, TESTI2046536

【 0 1 9 7 】

細胞骨格関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

BRACE3026993

【 0 1 9 8 】

核蛋白質・RNA合成関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

TESTI2046536

【 0 1 9 9 】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

TBAES2004105, UTERU2036507

【 0 2 0 0 】

発生、分化関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

BRCAN2019772

【 0 2 0 1 】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 8 クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI2046536, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 2 0 2 】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 2 クローンであった。

CTONG2003764, TRACH3021066

【 0 2 0 3 】

以下の 6 クローンについては、現在のところ相同性検索の情報からは機能を推定できる情報の得られないクローンであった。これらクローンについては今後、

データベースのアップデートによって機能が明らかになる可能性がある。クローン名を以下に示した。

BRAMY3008096, BRAMY3016953, BRHIP3038037, BRTHA3004432, TESTI4052089, UT ERU3020583

【0 2 0 4】

実施例 7. 推定アミノ酸配列に対する機能ドメインの検索による機能カテゴリー分類

ドメイン、モチーフはタンパク質の最小限の機能構造である。一タンパク質の構造はこの最小限構造の寄せ集めで成り立ち、その結果、タンパク質全体としての機能が決定される。よってドメインやモチーフ構造の解析から全体としてのタンパク質が持つ機能を比較的正確に予測することが可能である。また、この結果を機能別にデータベース化することは、特定の機能を持つクローンが容易に選択可能ということであり、個々のクロンの機能解析の際に非常に有用である。

全長塩基配列から推定されたアミノ酸配列のPfamに対するドメイン検索の結果（実施例 5 参照）から、ヒットデータのドメイン、モチーフ名やアクセッション番号、Pfam (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>) における詳細な記述データや、PROSITE (<http://www.expasy.ch/cgi-bin/prosite-list.pl>) にある機能カテゴリー分類を参照に、ヒットした 2 9 クローン中にコードされるタンパク質の機能予測、カテゴリー分類を行った。

【0 2 0 5】

分泌・膜蛋白質のカテゴリーに属すると推定されたクローンとは、受容体、イオンチャンネル、ホルモン、成長因子などと推測されるような例えば 7 transmembrane receptor, Pancreatic hormone peptides, Ion transport protein, Fibroblast growth factor等のドメイン、モチーフを持つクローンである。

【0 2 0 6】

糖蛋白質関連蛋白質のカテゴリーに属すると推定されたクローンとは、糖タンパク質、糖転移酵素などGlycobiologyに関わると推測されるような例えばImmunoglobulin domain, Glycosyl transferases group 1等のドメイン、モチーフを持つクローンである。

【0 2 0 7】

シグナル伝達関連蛋白質のカテゴリーに属すると推定されたクローンとは、タンパク質リン酸化酵素、脱リン酸化酵素、SH2ドメイン、Small Gタンパク質などと推測されるような例えばEukaryotic protein kinase domain, Protein phosphatase 2C, Ras family等のドメイン、モチーフを持つクローンである。

【0 2 0 8】

転写関連蛋白質のカテゴリーに属すると推定されたクローンとは、転写因子、転写調節に関わるタンパク質などと推測されるような例えばbZIP transcription factor, Zinc finger, C2H2 type等のドメイン、モチーフを持つクローンである。

【0 2 0 9】

疾患関連蛋白質のカテゴリーに属すると推定されたクローンとは、特定の疾患で発現が見られるようなタンパク質や、疾患で発現が上昇したり減少したりすると推測されるような例えばWilm's tumour protein, von Hippel-Lindau disease tumor suppressor protein等のドメイン、モチーフを持つクローンである。

【0 2 1 0】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンとは、転移酵素、合成酵素、加水分解酵素などと推測されるような例えばAldehyde dehydrogenase family, Chitin synthase, Glucose-6-phosphate dehydrogenase等のドメイン、モチーフを持つクローンである。

【0 2 1 1】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンとは、サイクリン、細胞増殖制御タンパク質などと推測されるような例えばCyclin, Cell division protein等のドメイン、モチーフを持つクローンである。

【0 2 1 2】

細胞骨格関連蛋白質のカテゴリーに属すると推定されたクローンとは、アクチン、キネシン、フィブロネクチンなどと推測されるような例えばActin, Fibronectin type I domain, Kinesin motor domain等のドメイン、モチーフを持つクローンである。

【0 2 1 3】

核蛋白質・RNA合成関連蛋白質のカテゴリーに属すると推定されたクローンとは、スプライシング因子、RNA合成酵素、ヘリカーゼなどと推測されるような例えばHepatitis C virus RNA dependent RNA polymerase, DEAD/DEAH box helicase等のドメイン、モチーフを持つクローンである。

【0 2 1 4】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンとは、翻訳関連タンパク質、ユビキチン関連タンパク質、Ribosomal proteinなどと推測されるような例えばTranslation initiation factor SUI1, Ubiquitin family, Ribosomal protein L16等のドメイン、モチーフを持つクローンである。

【0 2 1 5】

細胞防御関連蛋白質のカテゴリーに属すると推定されたクローンとは、分子シャペロン、DNA修復タンパク質などと推測されるような例えばHsp90 protein, DNA mismatch repair protein等のドメイン、モチーフを持つクローンである。

【0 2 1 6】

発生・分化関連蛋白質のカテゴリーに属すると推定されたクローンとは、器官形成関連タンパク質などと推測されるような例えばFloricaula / Leafy protein等のドメイン、モチーフを持つクローンである。

【0 2 1 7】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンとは、転写因子、DNAリガーゼをはじめとしたDNA・RNA関連酵素類、Zinc-finger関連タンパク質などと推測されるような例えばTranscription factor WhiB, B-box zinc finger, tRNA synthetases class I (C)等のドメイン、モチーフを持つクローンである。

【0 2 1 8】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンとは、ATPase等をはじめとしたATP・GTP関連酵素類、Gタンパク質などと推測されるような例えばE1-E2 ATPase, Ras family等のドメイン、モチーフを持つクローンである。

【0 2 1 9】

なお、この機能カテゴリー分類では一つのクローンが上記の複数のカテゴリーに該当する場合は、そのまま複数のカテゴリーに分類した。ただし、蛋白質の機能は必ずしも分類された機能カテゴリーに限定されるわけではない。

【0220】

分泌・膜蛋白質に属すると推定されたクローンは、以下の10クローンであった。

BRAWH3013197, BRAWH3028645, BRAWH3046240, BRCAN2019772, PROST2010326, TESTI2043585, TESTI4005158, THYMU3024602, THYMU3046350, TRACH1000193

【0221】

糖蛋白質関連蛋白質に属すると推定されたクローンは、以下の2クローンであった。

BRAWH3028645, TESTI4005158

【0222】

シグナル伝達関連蛋白質に属すると推定されたクローンは、以下の3クローンであった。

BRAMY3008096, TESTI4052089, TRACH3021066

【0223】

転写関連蛋白質に属すると推定されたクローンは、以下の8クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI2046536, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【0224】

酵素・代謝関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の7クローンであった。

ADRGL2011190, BRAMY3008096, CTONG2002832, TESTI4052089, THYMU3024602, THYMU3044175, TRACH3021066

【0225】

細胞分裂・増殖関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の1クローンであった。

TBAES2007481

【 0 2 2 6 】

蛋白質合成・輸送関連蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

CTONG2002832

【 0 2 2 7 】

DNA・RNA結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 7 クローンであった。

BRACE3002184, BRCAN2019772, CTONG2002832, TESTI4002988, TESTI4005500, TRACH1000193, TRACH3019290

【 0 2 2 8 】

ATP・GTP結合蛋白質のカテゴリーに属すると推定されたクローンは、以下の 1 クローンであった。

ADRGL2011190

【 0 2 2 9 】

以下の 7 クローンについては、Pfamでヒットデータ（実施例 5 参照）があったものの、上記のいずれのカテゴリーに属するか明らかでないクローンであった。今後同様のドメイン、モチーフを持つタンパク質のデータの蓄積と共に機能がより詳細に解明され、上記のカテゴリーに分類できる可能性がある。

BRACE3026993// TSC-22/dip/bun family

BRACE3046450// PDZ domain (Also known as DHR or GLGF). //PDZ domain (Also known as DHR or GLGF). //PDZ domain (Also known as DHR or GLGF). 1//PDZ domain (Also known as DHR or GLGF).

CTONG2003764// Phorbol esters/diacylglycerol binding domain (C1 domain)

TBAES2004105// Thrombospondin type 1 domain

TBAES2007379// EGF-like domain//EGF-like domain//Trypsin Inhibitor like cysteine rich domain//EGF-like domain//EGF-like domain//Keratin, high sulfur B2 protein//EGF-like domain//EGF-like domain//Granulins//EGF-like domain//EGF-like domain//EGF-like domain

TBAES2008133// PDZ domain (Also known as DHR or GLGF).

UTERU2036507// Thrombospondin type 1 domain//NTR/C345C module

【 0 2 3 0 】

またこれら以外にPfamでヒットデータがなかった残りのクローンについても、今後タンパク質のデータの蓄積と共に新たなドメイン、モチーフが見い出された場合、再びクローンの推定アミノ酸配列を新しいデータベースに対して解析することで新たな機能を有したドメイン、モチーフが発見され、カテゴリー分類できる可能性がある。

【 0 2 3 1 】

実施例 8. in silicoにおける発現頻度解析

実施例 1 に示した様々な組織・細胞由来のcDNAライブラリーを作製し、各ライブラリーからcDNAクローンを無作為に選択して、その 5' 末端領域の配列を決定し、データベース化した。本データベースは 1, 4 0 2, 0 7 0 個のクローンの塩基配列をデータベース化したものであり、解析母数としては十分なデータベースである。

次にこのデータベースにある各クローンの塩基配列を、塩基配列の相同性検索プログラムによって相同な配列同士をカテゴライズし（クラスター化）、各クラスターに属するクローン数を各ライブラリー毎に集計し規格化することによって、ある遺伝子のcDNAライブラリー内での存在比を解析した。この解析によって、cDNAライブラリーのソースとなっている組織や細胞における、ある遺伝子の発現頻度情報を得た。

【 0 2 3 2 】

次に本発明のcDNAの塩基配列を持つ遺伝子の、組織や細胞間での発現を解析するために、大量のcDNAクローンを解析した組織や細胞由来のライブラリーを組織・細胞間での発現量の比較の対象にした。すなわち 6 0 0 個以上のcDNAクローンの塩基配列を解析した組織や細胞について、先に規格化した数値を組織間や細胞間で比較し、遺伝子の発現頻度の変化を解析した。この解析によって以下に続く病態や機能に関連する遺伝子であることが示された。なお、以降に示される表 3 ～表 9、表 1 0、および表 1 1 ～表 1 3 中の各数値は、相対的な発現頻度を示し

、数値が大きいほど発現量が多いことを示す。なお、表の一部に比較したライブラリー間ではさほど大きな差がない遺伝子も含まれるが、他の組織、細胞由来のライブラリーと比較した場合は、有意な差が認められるので、それぞれの組織、細胞に特異的な遺伝子であり、疾患の診断マーカーや、分子メカニズム解明に有用な遺伝子と言える。

【0 2 3 3】

アルツハイマー病関連遺伝子

アルツハイマー病とは記憶力が低下し、進行すれば生活が困難となり介護が必要となる脳神経系の疾患であり、進行すれば脳そのものが萎縮する。その発症の要因はストレスなどの環境因子、高血圧やコレステロール血症などの血管因子も関わりがあるといわれているが、未だ不明である。したがって、正常脳組織とアルツハイマーの病態組織を比較した時、発現に差のある遺伝子はアルツハイマー病に関連する遺伝子であり、病態の発症メカニズムの解明や、遺伝子診断に有用であると考えられる。アルツハイマー患者の脳皮質由来のライブラリー（BRALZ、BRASW）と、正常全脳組織由来のライブラリー（BRAWH）のcDNAを解析して比較した結果（表3）、両者で発現変化のある遺伝子は以下の5クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子はアルツハイマー病に関する遺伝子である。

【0 2 3 4】

パーキンソン病関連遺伝子

パーキンソン病とは脳内の黒質で作られるドーパミンという神経伝達物質が十分量作られなくなり、その結果、手が震え、筋肉の動きが固くなって身体の動きが鈍くなる等の運動障害を引き起こす脳神経系の疾患である。脳の神経細胞は通常、歳を取るにつれて少しずつ減少するが、パーキンソン病では黒質の神経細胞が普通よりも早く著しく減少する。よって脳組織全体と黒質とを比較した時、発現に差のある遺伝子は黒質特異的な変動をするパーキンソン病に関連する遺伝子であり、発症メカニズムの解明や遺伝子診断に有用であると考えられる。黒質由来のライブラリー（BRSSN）と、正常全脳組織由来のライブラリー（BRAWH）のcD

NAを解析して比較した結果（表 4）、両者で発現変化のある遺伝子は以下の 5 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子はパーキンソン病に関する遺伝子である。

【 0 2 3 5 】

短期記憶・痴呆症に関連する遺伝子

脳組織の中で海馬とは記憶を扱う非常に重要な部位であり、得た情報の情報の要・不要を判断して、他の脳部位に記憶を蓄えさせる、記憶固定の働きがある。臨床的知見より、海馬に異常をきたしたり最悪海馬が無くなると、5分程度しか新しいことを覚えていられなくなる。また痴呆症患者の一部はこの海馬に異常をきたしていると考えられている。脳組織全体と海馬とを比較した時、発現に差のある遺伝子は記憶に関与したり、痴呆症に関連する遺伝子であり、記憶のメカニズム解明や遺伝子診断に有用であると考えられる。海馬由来のライブラリー（BRHIP）と、正常全脳組織由来のライブラリー（BRAWH）のcDNAを解析して比較した結果（表 5）、両者で発現変化のある遺伝子は以下の 5 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, TRACH1000193

これらの遺伝子は記憶および痴呆症に関する遺伝子である。

【 0 2 3 6 】

平衡感覚・運動機能に関する遺伝子

小脳は平衡感覚と筋肉運動、運動学習の中枢である。この領域は運動の調節に関与していると考えられており、小脳が動作することによって無意識的にスムーズな運動をすることが可能になる。また、運動だけでなく読み書きなどより高次の運動の慣れにも小脳が関与していることも最近の研究で解明されつつある。脳組織全体と小脳とを比較した時、発現に差のある遺伝子は平衡感覚や運動機能に関与する遺伝子であり、脳が制御する運動機能の分子メカニズム解明に有用であると考えられる。小脳由来のライブラリー（BRACE）と、正常全脳組織由来のライブラリー（BRAWH）のcDNAを解析して比較した結果（表 6）、両者で発現変化のある遺伝子は以下の 9 クローンであった。

BRACE3002184, BRACE3026993, BRACE3046450, BRAMY3008096, BRAWH3013197, BR

AWH3028645, BRTHA3004432, TESTI4002988, TRACH1000193

これらの遺伝子は平衡感覚および運動機能に関する遺伝子である。

【 0 2 3 7 】

感覚器からの情報伝達に關与する遺伝子

視床は、大脳と結びつきの強い神経細胞が集まった部分であり、脊髄などから伝わってきた感覚情報を大脳の関係部分に伝えたり、大脳の運動の指令を調節する。例えば視覚では映像を大きさ、形、色に分け、聴覚では音声を音量、耳障りの良し悪しで分け、大脳皮質の感覚野に送る。脳組織全体と視床とを比較した時、発現に差のある遺伝子は感覚器からの情報伝達に關与する遺伝子であり、脳が制御する情報伝達の分子メカニズム解明に有用であると考えられる。視床由来のライブラリー (BRTHA) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 7)、両者で発現変化のある遺伝子は以下の 8 クローンであった。

BRACE3046450, BRAWH3013197, BRAWH3028645, BRTHA3004432, BRTHA3024233, CTONG2002832, THYMU3046350, TRACH1000193

これらの遺伝子は感覚器からの情報伝達に關する遺伝子である。

【 0 2 3 8 】

情動反応に關する遺伝子

扁桃は脳の感情中枢である。扁桃を通過した情報は感情反応、例えばパニックや恐怖反応などを引き起こす。刺激が扁桃で情動評価されて強い恐怖を生じたとき、扁桃は脳の各部に警戒信号を出す。その結果、手の平の発汗、心悸亢進、血圧上昇、アドレナリンの急激分泌等の反応が起きる。いわば扁桃体は身体に警戒信号を送り、その結果として体を警戒態勢に入らせる一種の防衛本能を司っている組織とも言える。脳組織全体と扁桃とを比較した時、発現に差のある遺伝子は情動反応に關与する遺伝子であり、感情反応や恐怖反応、パニックなどの分子メカニズム解明に有用であると考えられる。扁桃由来のライブラリー (BRAMY) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果 (表 8)、両者で発現変化のある遺伝子は以下の 9 クローンであった。

BRACE3046450, BRAMY3008096, BRAWH3013197, BRAWH3028645, BRTHA3004432, TE

STI4002988, THYMU3046350, TRACH1000193, TRACH3019290

これらの遺伝子は情動反応に関する遺伝子である。

【 0 2 3 9 】

癌関連遺伝子

癌の組織では、正常組織とは異なる遺伝子のセットが発現して組織・細胞の癌化に寄与していると考えられている。したがって、正常組織とは異なる発現をする遺伝子は癌関連遺伝子である。正常な組織と比較して癌組織で発現変化する遺伝子を探索した。

【 0 2 4 0 】

乳がん由来のライブラリー (TBAES) と、正常な乳房由来のライブラリー (BEAST) のcDNAを解析して比較した結果 (表 9)、両者で発現変化のある遺伝子は以下の 4 クローンであった。

TBAES2004105, TBAES2007379, TBAES2007481, TBAES2008133

【 0 2 4 1 】

肝臓がん由来のライブラリー (TLIVE) と、正常な肝臓由来のライブラリー (LIVER) のcDNAを解析して比較した結果 (表 10)、両者で発現変化のある遺伝子は以下の 2 クローンであった。

THYMU3046350, TLIVE2007192

【 0 2 4 2 】

子宮がん由来のライブラリー (TUTER) と、正常な子宮由来のライブラリー (UTERU) のcDNAを解析して比較した結果 (表 11)、両者で発現変化のある遺伝子は以下の 3 クローンであった。

UTERU2017492, UTERU2025415, UTERU2036507

【 0 2 4 3 】

舌がん由来のライブラリー (CTONG) と、正常な舌由来のライブラリー (NTONG) のcDNAを解析して比較した結果 (表 12)、両者で発現変化のある遺伝子は以下の 2 クローンであった。

CTONG2002832, CTONG2003764

これらの遺伝子は、癌に関する遺伝子である。

【0244】

また、発生や分化に関連する遺伝子を調べる方法として、発生・分化途中の組織・細胞と、成体の組織細胞において遺伝子発現量の違いを調べる発現頻度解析がある。組織の発生・分化に関する遺伝子は、その組織の構築と機能発現に関する遺伝子であり、傷害のある組織を任意に再生せしめる再生医学に利用可能な有用な遺伝子である。

【0245】

先に記した1,402,070個のクローンの塩基配列のデータベースを基にして得た遺伝子発現頻度情報を用いて、発生・分化途中の組織・細胞と成体の組織・細胞とを比較して遺伝子発現頻度に変化のある遺伝子を解析した。

【0246】

胎児の脳由来のライブラリー（FCBBF, FEBRAまたはOCBBF）と成体の脳由来のライブラリー（BRACE, BRALZ, BRAMY, BRAWH, BRCAN, BRCOC, BRHIP, BRSSN, BRSTNまたはBRTHA）のcDNAを解析し、胎児と成体で比較した結果（表13）、両者で発現変化のある遺伝子は以下の13クローンであった。

BRACE3002184, BRACE3026993, BRACE3046450, BRAMY3008096, BRAWH3013197, BRAWH3028645, BRTHA3004432, BRTHA3024233, CTONG2002832, TESTI4002988, THYMU3046350, TRACH1000193, TRACH3019290

これらの遺伝子は組織・細胞の再生に関する遺伝子である。

【0247】

【表3】

アルツハイマー患者の脳皮質由来のライブラリー（BRALZ、BRASW）と、正常全脳組織由来のライブラリー（BRAWH）のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRALZ	BRASW
BRACE3046450	21.984	0.000	0.000
BRAWH3013197	100.000	0.000	0.000

BRAWH3028645	100.000	0.000	0.000
BRTHA3004432	32.087	0.000	0.000
TRACH1000193	21.722	0.000	0.000

【 0 2 4 8 】

【表 4】

黒質由来のライブラリー (BRSSN) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRSSN
BRACE3046450	21.984	0.000
BRAWH3013197	100.000	0.000
BRAWH3028645	100.000	0.000
BRTHA3004432	32.087	0.000
TRACH1000193	21.722	0.000

【 0 2 4 9 】

【表 5】

海馬由来のライブラリー (BRHIP) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRHIP
BRACE3046450	21.984	0.000
BRAWH3013197	100.000	0.000
BRAWH3028645	100.000	0.000
BRTHA3004432	32.087	0.000
TRACH1000193	21.722	0.000

【 0 2 5 0 】

【表 6】

小脳由来のライブラリー (BRACE) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRACE
BRACE3002184	0.000	19.020
BRACE3026993	0.000	100.000
BRACE3046450	21.984	78.016
BRAMY3008096	0.000	41.329
BRAWH3013197	100.000	0.000
BRAWH3028645	100.000	0.000
BRTHA3004432	32.087	0.000
TESTI4002988	0.000	15.392
TRACH1000193	21.722	15.417

【 0 2 5 1 】

【表 7】

視床由来のライブラリー (BRTHA) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRTHA
BRACE3046450	21.984	0.000
BRAWH3013197	100.000	0.000
BRAWH3028645	100.000	0.000
BRTHA3004432	32.087	35.583

BRTHA3024233	0.000	100.000
CTONG2002832	0.000	12.432
THYMU3046350	0.000	15.320
TRACH1000193	21.722	0.000

【 0 2 5 2 】

【表 8】

扁桃由来のライブラリー (BRAMY) と、正常全脳組織由来のライブラリー (BRAWH) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BRAWH	BRAMY
BRACE3046450	21.984	0.000
BRAMY3008096	0.000	58.671
BRAWH3013197	100.000	0.000
BRAWH3028645	100.000	0.000
BRTHA3004432	32.087	32.330
TESTI4002988	0.000	21.851
THYMU3046350	0.000	6.960
TRACH1000193	21.722	0.000
TRACH3019290	0.000	27.871

【 0 2 5 3 】

【表 9】

乳がん由来のライブラリー (TBAES) と、正常な乳房由来のライブラリー (BEAST) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	BEAST	TBAES
----------	-------	-------

TBAES2004105	0.000	100.000
TBAES2007379	0.000	100.000
TBAES2007481	0.000	100.000
TBAES2008133	0.000	100.000

【 0 2 5 4 】

【表 1 0】

肝臓がん由来のライブラリー (TLIVE) と、正常な肝臓由来のライブラリー (LIVER) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	LIVER	TLIVE
THYMU3046350	0.000	47.489
TLIVE2007192	0.000	100.000

【 0 2 5 5 】

【表 1 1】

子宮がん由来のライブラリー (TUTER) と、正常な子宮由来のライブラリー (UTERU) のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	UTERU	TUTER
UTERU2017492	5.197	0.000
UTERU2025415	100.000	0.000
UTERU2036507	100.000	0.000

【 0 2 5 6 】

【表 1 2】

舌がん由来のライブラリー (CTONG) と、正常な舌由来のライブラリー (NTONG)

のcDNAを解析して比較した結果、両者で発現変化のある遺伝子

Clone ID	NTONG	CTONG
----------	-------	-------

CTONG2002832	0.000	20.880
--------------	-------	--------

CTONG2003764	0.000	24.325
--------------	-------	--------

【 0 2 5 7 】

【表 1 3】

胎児の脳由来のライブラリー（FCBBF, FEBRAまたはOCBBF）と成体の脳由来のライブラリー（BRACE, BRALZ, BRAMY, BRAWH, BRCAN, BRCOC, BRHIP, BRSSN, BRSTNまたはBRTHA）のcDNAを解析し、胎児と成体で比較した結果、両者で発現変化のある遺伝子

Clone ID	FCBBF	FEBRA	OCBBF	BRACE	BRALZ	BRAMY	BRAWH	BRCAN	BRCOC	BRHIP	BRSSN	BRSTN	BRTHA
BRACE3002184	0.000	0.000	33.290	19.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BRACE3026993	0.000	0.000	0.000	100.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BRACE3046450	0.000	0.000	0.000	78.016	0.000	0.000	21.984	0.000	0.000	0.000	0.000	0.000	0.000
BRAMY3008096	0.000	0.000	0.000	41.329	0.000	58.671	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BRAWH3013197	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000	0.000	0.000	0.000	0.000	0.000
BRAWH3028645	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000	0.000	0.000	0.000	0.000	0.000
BRTHA3004432	0.000	0.000	0.000	0.000	0.000	32.330	32.087	0.000	0.000	0.000	0.000	0.000	35.583
BRTHA3024233	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000
CTONG2002832	0.000	0.000	27.853	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	12.432
TEST14002988	0.000	0.000	0.000	15.392	0.000	21.851	0.000	0.000	0.000	0.000	0.000	0.000	0.000
THYMU3046350	0.000	0.000	0.000	0.000	0.000	6.960	0.000	0.000	24.471	0.000	0.000	0.000	15.320
TRACH1000193	0.000	0.000	0.000	15.417	0.000	0.000	21.722	0.000	0.000	0.000	0.000	0.000	0.000
TRACH3019290	0.000	0.000	0.000	0.000	0.000	27.871	0.000	0.000	0.000	0.000	0.000	0.000	0.000

【0258】

【発明の効果】

本発明により、39のポリヌクレオチドが提供された。全長cDNAの分離が進んでいないヒトにおいて、新規な全長cDNAを提供した意義は大きい。分泌蛋白質、膜蛋白質、シグナル伝達関連蛋白質、糖蛋白質関連蛋白質、転写関連蛋白質等は、多くの疾患に関連した蛋白質であることがわかっている。疾患に関連した遺伝子や蛋白質は、診断マーカー、発現や活性を制御する医薬品の開発、あるいは遺伝子治療のターゲットになるなど医薬品の開発等に有効である。

中でも、分泌蛋白質をコードするcDNAは、蛋白質自身に医薬品としての有用性が期待できること、および多くの疾患に関連する遺伝子を含む可能性があることから、本発明によって提供されたこれらのcDNAは、産業上きわめて重要である。さらに、膜蛋白質、シグナル伝達関連蛋白質、転写関連蛋白質、あるいは疾患関連蛋白質といった蛋白質やそれをコードする遺伝子についても、疾患の指標となること等が期待できる。これらのcDNAも、産業上きわめて重要であり、コードする蛋白質の持つ活性の制御や、発現の制御を通じて疾患の治療効果をもたらすこと等が期待される。

なお本明細書において引用された全ての先行技術文献は、参照として本明細書に組み入れられる。

【0259】

相同性検索結果データ

全長塩基配列及び推定アミノ酸配列に対する相同性検索結果データを以下に示す。

各データは配列名、ヒットデータのDefinition、P値、比較配列の長さ、相同性、ヒットデータのAccession No.の順に//で区切って記載した。相同性検索の結果、既知のどの配列にもヒットしなかったものについてはクローン名のみ記載した。

ADRGL2011190// cGMP-dependent 3',5'-cyclic phosphodiesterase (EC 3.1.4.17) (Cyclic GMP stimulated phosphodiesterase) (CGS-PDE).// 0// 570aa// 97%// 000408

BRACE3002184// Zinc finger protein 85 (Zinc finger protein HPF4) (HTF1).
// 0// 435aa// 76%// Q03923

BRACE3026993// plectin 1, intermediate filament binding protein, 500kD [
Homo sapiens]// 4.00E-09// 107aa// 21%// NM_000445

BRACE3046450// Interleukin-16 precursor (IL-16) (Lymphocyte chemoattract
ant factor) (LCF).// 0// 530aa// 83%// Q14005

BRAMY3008096

BRAMY3016953

BRAWH3013197// Metabotropic glutamate receptor 4 precursor.// 0// 675aa/
/ 91%// Q14833

BRAWH3028645// CL1BA protein [Rattus norvegicus]// 2.00E-61// 126aa// 38
%// NM_022962

BRAWH3046240// Probable G protein-coupled receptor GPR35.// 1.00E-157//
273aa// 88%// Q9HC97

BRCAN2019772// Orphan nuclear receptor NR2E1 (Nuclear receptor TLX) (Tai
lless homolog) (T11) (hT11).// 0// 347aa// 92%// Q9Y466

BRHIP3030064// Homo sapiens renal tumor antigen (RAGE), mRNA// 8.00E-49/
/ 92aa// 98%// NM_014226

BRHIP3038037

BRTHA3004432

BRTHA3024233// Probable low-affinity copper uptake protein 2 (hCTR2) (Co
pper transporter 2).// 4.00E-70// 128aa// 90%// O15432

CTONG2002832// Zinc finger protein 264.// 0// 430aa// 72%// O43296

CTONG2003764// Protein kinase C, delta type (EC 2.7.1.-) (NPKC-delta).//
2.00E-91// 163aa// 91%// Q05655

PLACE7013963// Nik related kinase [Mus musculus]// 0// 447aa// 46%// NM_
013724

PROST2010326// G protein-coupled receptor 56; EGF-TM7-like [Homo sapiens
]// 0// 463aa// 94%// NM_005682

TBAES2004105// ADAM-TS 6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 6) (ADAMTS-6) (ADAM-TS6).// 1.00E-57// 110aa// 41%// Q9UKP5

TBAES2007379// CEGP1 protein [Homo sapiens]// 0// 545aa// 67%// NM_020974

TBAES2007481// Cyclin I.// 5.00E-51// 111aa// 44%// Q9Z2V9

TBAES2008133// semaF cytoplasmic domain associated protein 3; semaphorin cytoplasmic domain-associated protein 3A [Mus musculus]// 0// 422aa// 54%// NM_018884

TESTI2043585// Olfactory receptor 5V1 (Hs6M1-21).// 1.00E-79// 141aa// 47%// Q9UGF6

TESTI2046536// Sperm histone P2 precursor (Protamine P2) (HP2, HP3, and HP4) (Basic proteins HPS1, HPS2, HPI2, and HPI1).// 1.00E-10// 33aa// 64%// P04554

TESTI4002988// Zinc finger protein 208.// 0// 311aa// 52%// 043345

TESTI4005158// Amiloride-sensitive sodium channel delta-subunit (Epithelial Na⁺ channel delta subunit) (Delta ENaC) (Nonvoltage-gated sodium channel 1 delta subunit) (SCNED) (Delta NaCH).// 0// 592aa// 87%// P51172

TESTI4005500// Zinc finger protein 184.// 0// 340aa// 47%// Q99676

TESTI4052089

THYMU3024602// Cell surface glycoprotein EMR1 precursor (EMR1 hormone receptor).// 0// 795aa// 86%// Q14246

THYMU3044175// Putative adenosylhomocysteinase (EC 3.3.1.1) (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase).// 1.00E-161// 266aa// 93%// 043865

THYMU3046350// Sodium- and chloride-dependent betaine transporter (Na⁺/Cl⁻ betaine/GABA transporter) (BGT-1).// 0// 440aa// 88%// P48065

TLIVE2007192

TRACH1000193// Orphan nuclear receptor HMR (Early response protein NAK1)

(TR3 orphan receptor).// 0// 503aa// 84%// P22736
TRACH3019290// Zinc finger protein 93 (Zinc finger protein HTF34) (Fragment).// 0// 387aa// 73%// P35789
TRACH3021066// Receptor protein-tyrosine kinase ERBB-2 precursor (EC 2.7.1.112) (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).// 0// 360aa// 88%// P04626
UTERU2017492// Calcipressin 1 (Down syndrome critical region protein 1) (Myocyte-enriched calcineurin interacting protein 1) (MCIP1) (Adapt78).// 2.00E-83// 150aa// 88%// P53805
UTERU2025415
UTERU2036507// ADAM-TS 6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 6) (ADAMTS-6) (ADAM-TS6).// 6.00E-48// 104aa// 37%// Q9UKP5
UTERU3020583

【 0 2 6 0 】

【配列表】

SEQUENCE LISTING

<110> Research Association for Biotechnology

<120> Full length cDNA

<130> BTR-A0303

<160> 82

<170> PatentIn Ver. 3.1

<210> 1

<211> 3355

<212> DNA

<213> Homo sapiens

<400> 1

```
gacgcgctga ccgagcgcag cggccgggcc ggcgggcggg cgggcggctg cgagcatggt    60
cctgggtgctg caccacatcc tcatcgctgt tgtccaattc ctcaggcggg gccagcaggt    120
cttcctcaag ccggacgagc cgccgccgcc gccgcagcca tgcgccgaca gcctgcagga    180
cgcttgctg agtctgggct ctgtcatcga catttcaggc ctgcaacgtg ctgtcaagga    240
ggccctgtca gctgtgctcc cccgagtggg aactgtctac acctacctac tggatggtga    300
gtcccagctc gtgtgtgagg acccccaca tgagctgccc caggaggggg aagtcgggtt    360
caccgacgag gacgagcatg tgatccagca ctgcttcac tacaccagca ccgtgctcac    420
cagcaccctg gccttcaga aggaacagaa actcaagtgt gagtgccagg ctcttctcca    480
agtggcaaag aacctcttca cccacctgga tgacgtctct gtcctgctcc aggagatcat    540
cacggaggcc agaaacctca gcaacgcaga gatctgctct gtgttctgc tggatcagaa    600
tgagctgggtg gccaagggtg tcgacggggg cgtgggtggat gatgagagct atgagatccg    660
catcccggcc gatcagggca tcgcgggaca cgtggcgacc acgggccaga tcctgaacat    720
ccctgacgca tatgcccata cgcttttcta ccgcggcgctg gacgacagca ccggcttccg    780
cacgcgcaac atcctctgct tccccatcaa gaacgagaac caggagggtca tcggtgtggc    840
cgagctgggtg aacaagatca atgggccatg gttcagcaag ttcgacgagg acctggcgac    900
ggccttctcc atctactgcg gcatcagcat cgccattct ctctataca aaaaagtgaa    960
tgaggctcag tatcgcagcc acctggccaa tgagatgatg atgtaccaca tgaaggtctc   1020
cgacgatgag tataccaaac ttctccatga tgggatccag cctgtggctg ccattgactc   1080
caattttgca agtttcacct ataccctcg ttccctgccc gaggatgaca cgtccatggc   1140
catcctgagc atgctgcagg acatgaattt catcaacaac taaaaattg actgcccgcac   1200
cctggcccgg ttctgtttga tggatgaagaa gggctaccgg gatccccct accacaactg   1260
gatgcagcc ttttctgtct ccacttctg ctacctgctc tacaagaacc tggagctcac   1320
caactacctc gaggacatcg agatctttgc ctgttttatt tcctgcatgt gtcatgacct   1380
ggaccacaga ggcacaaaca actctttcca ggtggcctcg aaatctgtgc tggctgcgct   1440
```

ctacagctct gagggctccg tcatggagag gcaccacttt gctcaggcca tcgccatcct 1500
caacacccac ggctgcaaca tctttgatca tttctcccgg aaggactatc agcgcatgct 1560
ggatctgatg cgggacatca tcttggccac agacctggcc caccatctcc gcattctcaa 1620
ggacctccag aagatggctg aggtgggcta cgaccgaaac aacaagcagc accacagact 1680
tctcctctgc ctctcatga cctcctgtga cctctctgac cagaccaagg gctggaagac 1740
tacgagaaag atcgcgagag tgatctacaa agaattcttc tcccaggagg acctggagaa 1800
ggccatgggc aacaggccga tggagatgat ggaccgggag aaggcctata tccctgagct 1860
gcaaatacagc ttcattggagc acattgcatg gcccattctac aagctgttgc aggacctgtt 1920
cccaaagcg gcagagctgt acgagcgctg ggccccaac cgtgagcact ggaccaaggt 1980
gtcccacaag ttcacatcc gcggcctccc aagtaacaac tcgctggact tcctggatga 2040
ggagtacgag gtgcctgac tggatggcac tagggccccc atcaatggct gctgcagcct 2100
tgatgctgag tgatccctc cagggacact tccctgcca ggccacctc cacagccctc 2160
cactggtctg gccagatgca ctgggaacag agccacgggt cctgggtcct agaccaggac 2220
ttcctgtgtg accctggaca agtactacct tcctgggcct cagctttctc gtctgtataa 2280
tggaagcaag acttccaacc tcacggagac tttgtaattt gttctctgag agcacagggg 2340
tgaccaatga gcagtgggcc ctactctgca cctctgacca caccttggca agtctttccc 2400
aagccattct ttgtctgagc agcttgatgg tttctccttg cccatttct gccccaccag 2460
atctttgtct ctttcccttt gaggactccc accctttggg gtctccagga tcctcatgga 2520
aggggaaggt gagacatctg agtgagcaga gtgtggcatc ttggaaacag tccttagttc 2580
tgtgggagga ctagaaacag ccgcggggcg aaggccccct gaggaccact actatactga 2640
tggtgggatt gggacctggg ggatacaggg gccccaggaa gaagctgcca gaggggcagc 2700
tcagtgtctt gcagagaggg gccctgggga gaagcaggat gggattgatg ggcaggaggg 2760
atccccgcac tgggagacag gccaggtat gaatgagcca gccatgcttc ctctgcctg 2820
tgtgacgtg ggcgagtctc tccccctgtc tgggccaac agggagcggg taagacaatc 2880
catgctctaa gatccatttt agatcaatgt ctaaaatagc tctatcgctc tgcggagtcc 2940
cagcagaggc tatggaatgt ttctgcaacc ctaaggcaca gagagcccaa ccctgagtgt 3000
ctcagaggcc ccctgagtgt tctccttggc ctgagcccct taccattcc tgcagccagt 3060
gagagacctg gcctcagccc tggcagggtt ctctcttcaa ggccatatcc acctgtgccc 3120
tggggcttgg gagaccccat agggccggga ctcttgggtc agcccggcca ctggcttctc 3180

tctttttctc cgtttcattc tgtgtgcgtt gtgggggtggg ggaggggggtc cacctgcctt 3240
acctttctga gttgccttta gagagatgcg tttttctagg actctgtgca actgtcgtat 3300
atgggtcccgt gggctgaccg ctttgtacat gagaataaat ctatttcttt ctacc 3355

<210> 2

<211> 4006

<212> DNA

<213> Homo sapiens

<400> 2

actcagacac gcagctggac cggacagggc ggcttccggg tttggcgcgg cttttgtctc 60
ttgctgcagc cggagctccg gtctcgtctt cactgctgtt tgcctctgc tcctagaggc 120
ccagactctg tggctctgtg acctgcaggt attgggagat ccacagctaa gactccagga 180
ccctctggaa gcctagaaat gggaccattg acatttatgg atgtggccat agaattctct 240
ctggaggagt ggcagtgcct ggacacagca cagcagaatt tatataggaa tgtcatgtta 300
gagaattata gaaacctggg attcctgggt attgctgtgt ctaatttaga cctgatcacc 360
tgtctggagc aaggaaaaga gccctggaat atgaagagac acgagatggg agccaaaccc 420
ccagctatgt gttctcattt tgccaaagac cttaggccag agcaatatat aaaaaattct 480
ttccaacaag tgatactgag aagatatgga aaatgtggat atcagaaagg ctgtaaaagt 540
gtggatgagc ataagttgca caaaggaggc cacaagggac ttaaccagtg tgtgacaact 600
accagagca aaatagttca gtgtgacaaa tacgtgaaag tctttcataa atattcaaat 660
gcaaagagac ataagataag acatactgga aaaaatcctt tcaaatgtaa agaattgtggc 720
aaatcatttt gcatgctttc acaactaact caacatgaga taattcatac tggagaaaaa 780
ccctacaaat gtgaagaatg tggcaaagct tttaaaaagt cctcaaacct tacgaatcat 840
aagataattc atactggaga gaaaccttac aaatgtgaag aatgtggaaa agcttttaac 900
cagtcctcaa ctcttactag acataagata attcatactg gagagaaact ctacaaatgt 960
gaagaatgtg gcaaagcttt taaccgggtcc tcaaacctta ctaaataaa gatagttcat 1020
actggagaga aaccctacaa atgtgaagaa tgtggcaaag cttttaaaca gtcctcaaac 1080

cttactaatc acaagaaaat tcatactgga gagaaaccct acaaatgtgg agaattgtggc 1140
aaagccttta ccctatcttc acaccttact acacataaga ggattcatac tggtgagaaa 1200
ccctacaaat gtgaagaatg tggcaaagct tttagtgtat tttcaaccct tactaaacat 1260
aagataattc atactgaaga gaaaccctac aaatgtgaag aatgtggcaa agcttttaac 1320
cggtcctcac accttactaa tcataaggta attcatactg gagagaaacc ctacaaatgt 1380
gaagaatgtg gtaaagcctt taccaagtcc tcaactctta cttatcataa ggtaattcat 1440
actggaaaaga aaccctacaa atgtgaagaa tgtggtaaag ccttttagtat attctcaatc 1500
cttactaaac ataaagtaat tcatactgaa gacaaaccct acaaatgtga agaattgtggc 1560
aaaactttta actactcctc aaatttttact aatcataaaa aaattcatac tggagagaaa 1620
ccctataagt gtgaagaatg tggcaaaagc tttattctgt cctctcatct tactacacat 1680
aagataattc atactggaga gaaaccctac aaatgtaaag aatgtggcaa agcttttaac 1740
cagtcctcaa cccttatgaa acataagata attcatactg gagagaaacc ctacaaatgt 1800
gaagaatgtg gcaaagcctt taaccagtcc ccaaaccctta ctaaacataa gagaattcat 1860
accaaagaga aaccctacaa atgtaaataa tgtgtcaa at ccttttagtgt attttcaact 1920
tttattaaaa cacaagataa ttcatactgg agagaaatcc tacaaatgtg ataaatgtgg 1980
caaagccttt aactggctct taatccttat gaaacaaaag ataattcata ctagagagaa 2040
accctacaaa tgtaaagaat gtggcaaagc atttaactgg tccctaaatc ttactaaaca 2100
taaaaaaact taaactggag agaaactaca aagctggaag acatcacagt aattctgaca 2160
acacctcaaa cttttctaaa cataaaataa atcatactgg tgagggaactt taaaaatgcg 2220
aaaaatgtga caaatcttt aaatgattgt cacacttgat tgtaggtaag ataattctta 2280
ctagaaaaac tcttacaagt gtgaagaatg tggcaaaatt tttaatgaat actcacacct 2340
tattgccagg aaagcattta tacttgagaa aaattgtacc aatataaaga ctgtgaaaaa 2400
gccattaatg tctgctcaca tcttactcaa catcagagag tacttaaaaa aagtattata 2460
aatttaatta ctgtcaaaaa gtatttcaga aaatataagc ttttaaagta tttattttga 2520
agacaaacat tacaaataaa agagggttgt aataccttta tttgtatcac agatcttatt 2580
gtacacattt tgtactagag aaaaccctga agcagttact taaactttgt tcaacatcag 2640
ggaattcaga ttggagaaaa accctgcaaa tgtaataaat ttgaaaaaaa agttttccaa 2700
aactacaact tagacaacac cagagggttc atactgagat atatttttgc agatgcagta 2760
tatattttaa acatttaatc caaataaagt ctatgtaaat atcagagaat tcacagtgga 2820

aatatctaag gtagggccgg gcatggtggc tcacgcctgt agtgccagca ctttgggagg 2880
ctgaggtggg cagatcatga ggtcaacaga ttgagaccat cctggtcaac atggtaaaac 2940
cctgtctcta ctaaaaatac aaaaattagc tgggcgtggc ggcacgtgcc tgtggtctca 3000
gctacttggg acgctgaggc aggaaaatcc gttgaacca ggaggcagag gttgtagtga 3060
gccgggattg caccactgca ctccagccgg gtgacagagt gagactccat ctcaaaaaca 3120
aacaactaa ggtatggatg cttcagacat tacactatat caaagtgtg actgttaaaa 3180
atgatcccaa actaacggtg gataaattta tttgtacata actttaaaag gagtagaaga 3240
ttttttggaa gtttgaatt acattcaaag tatacttttt ttgtttgttt gttttccttg 3300
agacagagtc tctctctgtc gtccaggctg gagtgcagtg gcacagtctc ggctcgtgc 3360
agcctctgcc tcctgggttc aagcaattct cctgcctcag cctactgagt ggctgagatt 3420
acaggtgcac accaccacgt ccggtgact tttgtatttt tagtggggac ggggtttcac 3480
catgttggtc agggtagtct tgaactcctg acctcatgct ctgcccagct tggcttccca 3540
aggtgctgtg attacaggct taagccacag cgcccgaccc aaagtgttct tttttccttg 3600
aaaagaagta cagatttttt aaaaagctaa tattgctgtg atjcaactct caagttactt 3660
gatgctattt cttcatttct attgtattca catgtgaaag catgtgatca gttgtcacgg 3720
catcaaagat aggagagttt cttttttttt aggtgggcat tatttatgac cttttctttg 3780
gaagagtaag gatattaaaa tgtaacatgc atgatgaaaa tctaagtgga gatgatattt 3840
gtggttaatt atgatattga gtggtgcatg aggtaggtgt ttagagtaat attcttttgc 3900
attatagtaa gaacatttaa gttttgttaa aattaaaaaa ttagtagtac gttattttac 3960
tatttgtact tttatgtaat aaaatgcagt acatttttaa attttc 4006

<210> 3

<211> 5392

<212> DNA

<213> Homo sapiens

<400> 3

gagtgttttc agaagtggga atagcaggcg aggtcttgag gcctgaagaa gcaaagcgtc 60

gcttagtcaa tgctgagatg cgtcttgtgg aacagaattt ttgggacatt tgcaggatga 120
gtagatgaa ctccgtgctg agatggaaga gatgagagac agttatttag aggaagatgt 180
ttaccagctg caggaacttc ggcgagaact ggaccgctgct aataaaaaact gccgaatcct 240
gcagtaccgt cttcggaaag ccgagcagaa aagcctgaaa gtggctgaga cgggtcaggt 300
ggatggtgag cttattcgaa gcctggagca ggacttgaag gtagccaaag atgtatctgt 360
cagattgcac cacgaactta agacgggtgga ggaaaagcgc gctaaagctg aggatgaaaa 420
cgaaactctc cgacagcaga tgattgaagt ggaaatatcc aaacaggccc tccagaatga 480
gctggagaga ctgaaagaga gtccctaaa aagaagaagc actcgggaaa tgtacaagga 540
gaagaaaacc tttaaccagg atgacagtgc cgatttgagg tgccagctcc agtttgccaa 600
agaggaagcc ttcctgatgc gcaaaaagat ggccaagcta ggaagggaga aggacgagct 660
ggagcaggag ctccagaagt acaagtcctt ctatggggat gtggacagtc ccctgcccac 720
gggggaagca ggcgggcccc ccagcacccg ggaggccgag ctgaagctgc ggctaaagct 780
ggtggaggag gaagccaaca tcttgggccg gaagatcgtg gagctggagg tggagaaccg 840
tggcctcaag gcagagatgg aggacacgcg gggccagcag gagcgggagg gcccgggtcg 900
ggaccacgca cccagcattc ctacctacc cttcgggtgac tccctggagt cctccactga 960
gctccgccgc cacctgcagt ttgtagaaga ggaagcggag ttgctccgga ggtccatctc 1020
cgagatcgaa gaccacaacc ggcaactgac ccacgagctc agcaagttta agtttgagcc 1080
tccccgggag ccgggctggc taggagaggg tgcaagtcct ggtgccgggg gtggggcccc 1140
cctgcaggag gagctgaagt cagccaggct gcagatcagc gagctcagcg gcaaggtgct 1200
caaaactgcag cacgagaacc acgcgctgct gtccaacatc cagcgtgctg acctggcagc 1260
ccacctgggg ctgcgtgccc ccagtccccg ggacagcgat gccgagagtg atgcgggcaa 1320
gaaggagagt gatggggagg agagccgcct gccccagccc aagcgggaag ggcctgttgg 1380
cggggagagt gactcggagg agatgtttga gaagacgtcg ggcttcggga gcgggaagcc 1440
atcggaggcc agcgagccat gccccacgga gctcctgaag gcccgggagg actctgagta 1500
cctagtgacc ctaaaacacg aggcccagcg gctagagcgg acggtggagc gcctcatcac 1560
ggacaccgac agcttcctcc atgatgcggg gctgcgggggt ggtgcgccct taccggggcc 1620
tggcctccag ggcgaagagg agcagggtga gggggaccag caggagcccc agctgctggg 1680
gaccatcaac gccaagatga aggctttcaa gaaagagctg caggccttcc tggagcaggt 1740
gaaccgcatt ggggggtggc tatccccctt gccccacctc acagagtcct ctagcttctt 1800

ctccactgtg acttccgtgt cccgggactc ccccatcggg aacctgggga aggagctggg 1860
cccagacttg cagtccagac tgaaagagca gctggagtgg cagctcgggc cggcccagg 1920
ggacgagcgg gagagcctgc gcctccgagc cgcgcgggag ctgcaccgcc gcgcagacgg 1980
ggacaccggg agccacgggc tgggaggcca gacctgcttc agcctggagc tcaggggccc 2040
ccccgtttta cctgagcaga gtgtatccat agaggagcta cagggtcagc tcgtgcaggc 2100
ggccagactg catcaagagg agacagagac atttacaac aagatccata agatggagga 2160
ggagcacctc tatgccttga ggtggaaaga actggaaatg cacagcctgg ctttgcaaaa 2220
caccctccat gagcgaacct ggagtgatga gaagaatctg atgcagcagg agctccggtc 2280
cttgaagcag aacattttcc tcttctacgt caaactcagg tggctgctga aacactggcg 2340
gcaagggaag cagatggagg aggaaggaga ggagttcact gaggggtgaac atccagagac 2400
cctctccagg ctcggggagc ttggagtcca ggggggtcac caggcggatg gcccagacca 2460
cgacagtga c gaggctgtg gctttccagt gggggagcac tccccacact cccgggtgca 2520
gattggagat cacagcttgc ggctgcagac cgcggacagg ggacagcccc acaaacaggt 2580
ggtggaaaac cagcagctgt tcagcgcctt caaggccttg ctggaggact tccgtgcgga 2640
gctgcgggag gatgagcgtg cccgactacg gctgcagcag caatatgcca gcgacaaggc 2700
ggcctgggac gtggagtggg ccgtgctcaa gtgccgtctg gaacagctgg aagagaagac 2760
tgagaacaag ttgggagaac taggctcctc cgctgagagc aagggggcct tgaagaagga 2820
gagagagggt caccagaagc tcctggcaga cagtcacagc ctggtcatgg acctgcgctg 2880
gcagatccat cacagcgaga agaactggaa ccgggagaag gtggaacttc tcgaccgcct 2940
ggacagagat cggcaggagt gggagcggca gaagaaggaa ttcttgtgga ggatagagca 3000
gttgcagaaa gagaacagtc cccggagagg tggcagtttc ctctgtgatc aaaaagacgg 3060
caacgttcgc cctttcccc accagggaag cctccgcatg ccccgccag tggccatgtg 3120
gccttgtgca gatgtgact ccatcccgtt tgaagaccgg ccgctgtcca agctgaagga 3180
gtcggacagg tgctcggcca gtgagaatct ctacctggat gccttgtccc tggatgacga 3240
gccagaagag ccaccagccc acaggcccga gagggagttc aggaaccgcc tccctgagga 3300
agaagaaaaat cacaaaggaa atcttcaaag ggcggtgtcc gtgtcctcca tgtctgagtt 3360
ccagcgtcta atggacatct ccccttccct gcctgagaag ggccgtccgt ccaccagcag 3420
caaggaggat gtcacccac ccctgtctcc agacgacctc aagtacatcg aggagttaa 3480
caagagctgg gactacacac ccaacagggg ccacaatggt ggggggccgg acctttgggc 3540

cgacaggacc gagtggggc gggcaggga cgaggacagc acagagcctt tccccgactc 3600
ctcctggtac ctaaccāaa gtgtcaccat gaccacggac accatgacca gcccagagca 3660
ctgccagaag cagccactgc ggagccacgt cctcaccgag cagtcggggt tgcgcgtgtt 3720
acacagcccg cctgccgtgc gcagggtcga cagcatcacg gcggcagggtg gtgagggtcc 3780
ctttcccaca agcagagcca gagggagccc gggagacacc aaggggggccc ctccagaacc 3840
catgctcagc aggtggcctt gcacctcccc caggcactcc cgggactatg tggagggggc 3900
acggcgcccc ctgatagtc ccctctgtac ctccctgggg tttgcctccc cactgcacag 3960
cctggagatg tccaagaact tgagtgatga catgaaggag gtggccttct ctgtcaggaa 4020
tgccatctgc tctggccctg gcgagctgca agtcaaggac atggcctgcc agaccaatgg 4080
gtcccggacg atggggaccc agactgttca gaccatcagt gtgggcttgc agactgaagc 4140
cctgcgtggc agcgggtgtca ccagcagccc ccacaagtgt ctactccaa aggctggggg 4200
cgggtgctaca cccgtgtcgt ctcttcccc gagccttagg agcagacagg tggcccctgc 4260
catcgagaag gtgcaggcca agtttgaacg cacatgctgc tcccccaagt atggttctcc 4320
caagctgcag aggaagcccc tccccaaagc cgaccagcca aataacagga cgtcaccagg 4380
gatggcccag aaagggtaca gtgagtcagc ctgggcccgc tccaccacca caagggagag 4440
ccccgtgcac accaccatta atgatggcct ctccagcctc ttcaacatca ttgaccacag 4500
ccccgtggtg caggaccct tccagaaggg gctgcgggccc ggcagtcggt ctgcctcagc 4560
agagccccga ccagagctgg gcccaggcca ggaaacaggc accaattccc gaggaaggtc 4620
gcctagcccc attgggggtg ggtcagagat gtgcaggag gaagggggag agggcacgcc 4680
agtgaagcag gacttatctg ctccccctgg ctacacgctc actgagaacg tggcccggat 4740
cctcaacaag aagctgctgg aacatgcctt aaaggaggag aggaggcagg ctgcccacgg 4800
gccccgggt ctccacagt acagccactc gctgggggac acagccgagc cagggcccat 4860
ggagaaccaa actgtcttgc taactgcccc ctggggactc tagccctgcc cgcctcacgc 4920
tgaactacct tgttctgcac tagctccatc cctagagccc tgcttctcca ggcccagag 4980
accagcaaac cgtcgccctc cgtcccgttg ggccccacat tccccactg cctcacagcc 5040
tcagtcaccc ggagaccga cgtccttgga ggagcatggt ggcgaggagc cgcccagga 5100
gcagccacac cgagatgcaa gcttgcatgg attatcacag tataattcac tgtaatttgc 5160
ataaccacac catcaccatg aacaaaactc tgcccaacag gagagatcta gttttctcaa 5220
ggtcaaagaa tgttttttaa aaacacaaag ctgctgaatg ttcaacctgt gaaactgaga 5280

tgtttctaga atgaaacagt aaatgtgcct gtaataactt aatTTTTTtc atagctcaga 5340
aaactatTTTt tgtctccatc cTTTTTtacac acagtatatatt aaacgaaaag gc 5392

<210> 4

<211> 5154

<212> DNA

<213> Homo sapiens

<400> 4

acatttgtcc tgagtcacct gtccagagca ggtggtgaat attgtgtcct actcacggca 60
tctcaactat cggagcctgg gatctgactc aaaggccggc ctccgtctga gaactgagcg 120
tccatttctc aatccttgcc ggctctgacc caggcctggg ccacaggctg tccgggaata 180
agtgggtgctg caatccctgc tgggcagatg gagagaggag caagggagat ggcagccccg 240
ggggactgtg catagggagg gactcatgaa gtggcagcta agccctgtcc agtggccacc 300
cgtcagccaa gggccagaga ccaggaaagg aagaaaggca gcttacttc ctctttgagg 360
atggagtcgc acagccgcgc tggaaagagc agaaaatctg caaaatttcg gtccatctcc 420
aggtccttga tgctctgtaa tgctaagacc agtgatgatg gctctagccc tgatgagaaa 480
tattctgatc cttttgagat ttctttggcc cagggcaagg agggaatttt ccaactcatct 540
gtgcagctgg cagacacatc ggaggctggg cccagcagtg ttcttgatct agcactggcc 600
tcggaggctg ctcaactcca agcagctggg aatgatcgag gcaagacctg taggaggata 660
ttcttcatga aggaatcttc cacagcttcc tctcgagaaa agcctggaaa actagaagca 720
caaagtagta acttcctgtt tcctaaagcc tgccaccaa gggcacgcag caactcaacc 780
agtgttaatc cctattgcac aagagaaatt gattttccaa tgaccaagaa atctgcagcg 840
cccacggaca ggcagcctta ctctctctgc agtaacagga agtccctctc tcaacaattg 900
gactgtccag caggaaaggc tgcgggaact tcgagaccaa cacggtccct gagcacagct 960
cagctcgtgc agccatctgg gggcctccag gcttcagtca tctccaacat cgtgctgatg 1020
aagggccagg ctaagggtct gggcttcagc atcgttgggg gaaaagacag catttatggc 1080
cccattggga ttacgtcaa aaccattttt gcagggggag cagcagcagc cgatggaagg 1140

ctacaggaag gtgatgaaat tctggagctc aatggatgaat caatggctgg actaacacat 1200
caggatgctt tgcagaagtt caagcaagcc aaaaaggggc tcctcaccct caccgtgaga 1260
acccgcctga cggcgcctcc ttccctgtgc agccacctgt ctccccact gtgccgctcc 1320
ctgagctcca gcacttgtat caccaaggac agcagctcct tcgccttggga aagccccctcg 1380
gtcccatca gcaccgcaa gcccaattac agaatcatgg tggaggtttc tctgcagaaa 1440
gaggccggcg tgggcctggg catcggcctg tgcagcgctc cctacttcca atgcatctct 1500
ggcattttcg tccacacgct gtcaccagga tccgtggcgc acctggacgg acgtctccgg 1560
tgtggggacg agatttgtga aatcagtgat tcccctgtgc actgcctgac gctcaatgaa 1620
gtctacacga tcctgagtca ctgtgatccc ggtccagtc ccatcattgt tagccgacat 1680
ccagaccac aggtctctga acagcaactc aaagaagctg tggcccaggc tgtggaaaac 1740
accaagtttg gaaaggagag gcatcaatgg agtctggaag gtgtcaaaag gctggaaagc 1800
agtggcacg ggcggccac cttggagaag gaacgagaga agaactcagc acccccgcat 1860
cgcagggtc agaaggtcat gatccgctcc agcagtgaca gcagctacat gtctgggtcc 1920
ccagggggaa gtcctgggag tggcagtgct gagaagccgt cctctgacgt ggacatcagc 1980
acacacagcc ccagcttgcc tctggcacgg gagccagtgg tgctttctat agcatcctcc 2040
aggctgcccc aggagagccc acccctccca gagagccggg acagccacc gccgctgaga 2100
ctgaagaaat cctttgagat tttggtgaga aagcctatgt cctccaagcc caagcctcca 2160
cccagaaaat actttaaaag tgacagtgac cctcagaaga gtctggaaga gagagagaac 2220
tcctcatgct cttctgggca caccaccacc acctgtggcc aggaagcgag agagctgctg 2280
ccactgctgc taccacagga agacacagca gggagaagcc ctagtgctc tgccggctgc 2340
ccaggacctg gtatcgccc acagaccaag tcctccacag agggcgagcc aggggtggaga 2400
agagccagcc cagtaccca aacatccccg ataaaacacc cactgcttaa gaggcaggct 2460
cggatggact atagctttga taccacagcc gaagaccctt gggttaggat ttctgactgc 2520
atcaaaaact tatttagccc catcatgagt gagaacctg gccacatgcc tctacagccc 2580
aatgccagcc tgaatgaaga agaagggaca cagggccacc cagatgggac cccaccaaag 2640
ctggacaccg ccaatggcac tcccaaagtt tacaagtcag cagacagcag cactgtgaag 2700
aaaggtcctc ctgtggctcc caagccagcc tggtttcgcc aaagcttgaa aggtttgagg 2760
aatcgtgctt cagaccaag agggctccct gatcctgcct tgtccacca gccagcacct 2820
gcttccaggg agcacctagg atcacacatc cgggcctcct cctcctctc ctccatcagg 2880

cagagaatca gctcctttga aacctttggc tcctctcaac tgcctgacaa aggagcccag 2940
agactgagcc tccagccctc ctctggggag gcagcaaac ctcttgggaa gcatgaggaa 3000
ggacggtttt ctggactctt ggggagggg gctgcacca ctcttgtgcc ccagcagcct 3060
gagcaagtac tgtcctcggg gtccccctgca gcctccgagg ccagagaccc aggtgtgtct 3120
gagtccccctc ccccagggcg gcagcccaat cagaaaactc tccccctgg cccggaccg 3180
ctcctaaggc tgctgtcaac acaggctgag gaatctcaag gccagtgct caagatgcct 3240
agccagcgag cacggagctt cccctgacc aggtcccagt cctgtgagac gaagctactt 3300
gacgaaaaga ccagcaact ctattctatc agcagccaag tgtcatcggc tgtcatgaaa 3360
tccttgctgt gccttccatc ttctatctcc tgtgccaga ctccctgcat cccaaggaa 3420
ggggcatctc caacatcatc atccaacgaa gactcagctg caaatgggtc tgctgaaaca 3480
tctgccttgg acacagggtt ctgctcaac ctttcagagc tgagagaata tacagagggt 3540
ctcacggaag ccaaggaaga cgatgatggg gaccacagtt cccttcagtc tggtcagtcc 3600
gttatctccc tgctgagctc agaagaatta aaaaaactca tcgaggaggt gaaggttctg 3660
gatgaagcaa cattaaagca attagacggt atccatgtca ccatcttaca caaggaggaa 3720
ggtgctggtc ttgggttcag cttggcagga ggagcagatc tagaaaacaa ggtgattacg 3780
gttcacagag tgtttccaaa tgggctggcc tcccaggaag ggactattca gaagggaat 3840
gaggttcttt ccatcaacgg caagtctctc aaggggacca cgcacatga tgccttggcc 3900
atcctccgcc aagctcgaga gccaggcaa gctgtgattg tcacaaggaa gctgactcca 3960
gaggccatgc ccgacctcaa ctctccact gactctgcag cctcagctc tgcagccagt 4020
gatgtttctg tagaatctac agaggccaca gtctgcacgg tgacactgga gaagatgtcg 4080
gcagggtgg gcttcagcct ggaaggaggg aagggtccc tacacggaga caagctctc 4140
accattaaca ggattttcaa aggagcagcc tcagaacaaa gtgagacagt ccagcctgga 4200
gatgaaatct tgcagctggg tggcactgcc atgcagggcc tcacacggtt tgaagcctgg 4260
aacatcatca aggcactgcc tgatggacct gtcacgattg tcatcaggag aaaaagcctc 4320
cagtccaagg aaaccacagc tgctggagac tcctaggcag gacatgctga agccaaagcc 4380
aataacacac agctaacaca cagctcccat aaccgtgat tctcagggtc tctgtgccg 4440
ccccaccag atgggggaaa gcacaggtgg gcttcccagt ggctgtgcc caggcccaga 4500
ccttctagga cgccaccag caaaagggtt ttctaaagt aagggcagag tcacacgggg 4560
gcagctgata caaattgcag actgtgtaaa aagagagctt aatgataata ttgtggtgcc 4620

acaaataaaa tggatttatt agaatttcat atgacattca tgcctggctt cgcaaaatgt 4680
 ttcaagtact gtaactgtgt catgattcac ccccaaacag tgacatttat ttttctcatg 4740
 aatctgcaat gtgggcagag attggaatgg gcagctcatc tctgtccac ttggcatcag 4800
 ctggcgatcat gcaaagtcac gcaaaggctg ggaccacgtg agatcattca ctcatacatc 4860
 tggccgttga tgttggctgg gaactcacct ggggctgctg gcctgaatgc ttataggtgg 4920
 cctctccttg tggcctgggc ctctcaca catggtgtct ggattcccag gatgagcatc 4980
 ccaggatcgc aagagccacg tagaagctgc atcttgttta tacctttgcc ttggaagttg 5040
 catggcatca cctccacat actccatcag ttagagctga cacaaacctg cctgggttta 5100
 aggggagagg aaatattgct ggggtcattt atgaaaaata cagtttgtca catg 5154

<210> 5

<211> 3539

<212> DNA

<213> Homo sapiens

<400> 5

gattttttca tttttccggg ttgtgccgat cccttgggag tggccagaca ggggtggtggt 60
 gactgaagtt atgaggacag gaactgggcc atgtagcaga tactgatggt tgtattgatg 120
 actttttaaa gtgaggtgca catctggggt catgacatat agtcatcata agtgtgtgcc 180
 tgacttttct ttaaatttct ttatatcggg ctccctaaaa acagtgttca cactagttaa 240
 cccaccaccc aaaccacaga cactgatagt gctggacatc tctcgggtgggt gtgccagttt 300
 gtaaggctgg gccacttta gttgggtcaa gaggagggca gagggcattc taagctgagg 360
 ggtgggaggg agcagaggtg aggggctgaa ctaggccagg acaccagag gcggaaagtc 420
 tgcgcccagg acagaaggga ggcaggagtg ggagccggcc ctggcatcca tactcttttt 480
 ctccctgtcc ccacgtcac cgctccctct gcatggccca ggaggcactc tctggtccca 540
 cctgctctcc caggcgact cccgacattc tgggctcagc tctggctcta cccaggagag 600
 gatgaaggct cagctcaacc cccacctcaa cctcctgacc ccagcgaggc ttccctcagg 660
 ccatgcccct ggccaggaca gaatcgccct ggagcctcct aggacttctc cgaaccttct 720

cctagctggg gagggcccat ccaccagacc ccagagggag gctgaagtg aaccacagc 780
caggaccagc acctctggct cctcggacct tccaaaggcc ccaggtggcc acctgcacct 840
tcaagctagg agggctggcc agaactcaga cgctgggccc cctcgggggc tcaattgggt 900
tcctgagggg gccggcccgg tgctaggggg ctgtggccga ggcatggatc agagctgcct 960
gtcagcagat ggggcccggc ggggctgtgg cagggccacc tggagtgtga gagaggagca 1020
ggtgaagcag tgggcggcag agatgctggt agcgctggag gcgctgcacg agcagggggt 1080
gctgtgccgg gacctccacc ccgggaacct gctcctggac caggcaggta ggtgcctcc 1140
tcacccctca gtggggtggc ccccaaagcg tccaggcaga ggacgagaga tcccaaagt 1200
cgaggggacc tccaaggttg tctgtcctc cctggcccca gccagccttt cagggggcct 1260
aggctggagt ctgttgccg aggtccttg gagtacctg tgggctctgt gggcctcaga 1320
gtgctgtgcc tctgtggagg gtgctgtgcc cctgccgagt ttctgcccag caaggccatt 1380
ggctcctacc gcatgactgt tgggggaagc cctgcaacc cctccctgt atcagctgtc 1440
ctcaatctgt cccctcctt ttctccac acactggagg tcacatccgg ctacatatt 1500
ttggccagt gtcagagtg gagccccagt gctgcgggga ggccgtggac aatctctaca 1560
gcgccccaga ggtgggtggg atttccgagc tgacggaagc ctgtgactgg tggagctttg 1620
ggtctctact gtatgaactg ctgacgggaa tggcactgtc ccagagccac ccttcaggaa 1680
tccaggccca caccagctc cagctgccc agtggctcag tcgcccagcg gcctctctgc 1740
tgactgaggt gaggtgtagc ccggccagag gaagcagctg cctagcctga ggaaggtgct 1800
gggaggggca cagaactgtg tctgaaccac gtgtgcttcg tgctcagtgt ctgatgtggg 1860
tgcagagggc agtaggggca acacatttct atatgttggg ggccaaccac caccaatgtt 1920
cccagggtaa aatgggggtt ctgtgggcca ggctgccagg gagcaggccc cctagaggcc 1980
atagtgggtg ggtccctgg aggtcaccac atgcctccac agctggagtg gccagctga 2040
ctgggcctct ctctgcagt gctgcagttc gagcctacc ggcgctggg catgggagaa 2100
ggtgggtgtca gcaaactcaa gtccatccc tttttcagta ccatccaatg gagcaagctg 2160
gtggggtaag agggcagagc gggtgacgga agcagctggc ctggtctgga tcgcctctcc 2220
tccttgccctg acaccaacc cagggtggc cctctatcag cgggctttgg gcgaggaatg 2280
gagggcactg tctgtcctgc tgggtccca ctggggcctc agaattatgg ccaccacca 2340
ggaagggcca gtcctggaa aagctggagg tgggggcagt caaggcttgc cctgctaagc 2400
agcttgaacc atctacccat cagtcaacag accggttgag catgtggact caccatgtta 2460

aaggttgcc tctgtggcac tggcgctgag ctgttgacca cctgctgcac cctactgtga 2520
ggttctgtga ctactcact gccatgttgt gcccactca ggacatctct ggagactcat 2580
ctcaggacac tgatccactg gctcagtga cccaaaccag actgtcctgg ctagtcctct 2640
tagtcacaca gcgagtaggc ctcttcacc agaagctgtt cgtgcagctg gggaggccca 2700
gaagcaggga tgccatctgt tgtgcaggga gggatgaggt aggggacagg atgaccccca 2760
ccccacccc cccaccccg attcctgacc atgccccta tcttgtgatt tcagacttgg 2820
gagagcaaaa gggaggggaa gaaaataaga gtaatggggg gaggaaggaa tgcattgggtc 2880
tgcccttag agcaagtctg aaaccagaat caagagtcct cctccccagt gggcttgtgt 2940
gggggaagag gggcagcctg cccaggggt gcaagaggac agcaattcag cttccaggcc 3000
aaagcagttt agacaagtg tgccatcag gactccctct caggccctgg ctgtttacca 3060
tgtcactgcc tactgtgact tcatgccctt ttgggataag aaaccacaaa acgcatggaa 3120
caactatcca tggctcttag ccagaaatgt cttgcccctg gtctctgaaa agctgggggtg 3180
agtggttacc tcaaagacaa ttgcctccag tttgggggct cctctcaggg ttcacctgt 3240
caccggggcg tgggaccgac ctgggaaatt tccactagag ttttgccctc agtcctccc 3300
tgccagcatc ccagtcacca ccaggtggcg cttctgacac ggtccacgtg cctggctctg 3360
aaggcccaac tgcagacca gaccctcac tcatggtggg gagacttgac ttctcgggc 3420
tcttcaggag actgagacag agaagacaag tgaaagcgca gcctgtacat gactctctgt 3480
ggtccacgtc ctgcctctcc ctggggttta cacggcccca ataaagcctg ctctggcct 3539

<210> 6

<211> 5680

<212> DNA

<213> Homo sapiens

<400> 6

agtacgcggt cgctactgat catgcttggg ccagggtcca atcgcaggcg cccacgcag 60
ggggagcgag gcccagggtc ccccgagag cccatggaga agtaccaggt tttgtaccag 120
ctgaatcctg gggccttggg ggtgaacctg gtggtggagg aaatggaaac caagtcaag 180

catgtgataa agcaggtgga atgcatggat gaccattacg ccagtcaggc cctggaggag 240
ctgatgccac tgctgaagct gcggcacgcc cacatctctg tgtaccagga gctgttcac 300
acgtggaatg gggagatctc ttctctgtac ctctgcctgg tgatggagtt caatgagctc 360
agcttccagg aggtcattga ggataagagg aaggcaaaga aaatcattga ctctgagatc 420
aagacagaag gcgttcctaa cactagaagg ttcccagtcg gtgacatga ttccagattg 480
ttctgccggt ccctgaagtt cctgtaaatt gactcgtccg gcatgctgcc actcctgtct 540
ggcttccttc cctcagcctg ctgttgtgag ccccgcggtg ctgctgcatg caccagcaaa 600
tcatgtgttc attgcttgct gccactctgc tgctgattg cgctgcaggc tgtttaccta 660
gtctcatttg ggctgcttc agtttggggc tatttgtaat aaggctgcta tgagcattgc 720
tggaagacac tcacttttct gggatgaata cctaggagtg gaattattgg gtcgtagagt 780
acatgtgtgt agcttcagtg gatgctccaa acagatttcc gacttggttt ggttggcccc 840
atgtttactc tcacaagttg tgagcattcc cgatccacat ggaggccagc acttcattgt 900
gtcagtcttg ttgttgtgt tgttgttgag atggagtctc tctctgtcgc ccaggctgga 960
gtgcagtggc acgaccttg ctcactgcag ccttcgcctc cctgttcgag cgattctcct 1020
gcctcagcct cccaagtagc tgggactaca ggtgcccacc accacacca ctaatttttg 1080
tattattaat agagacaggg ttttgctatg ttgccaggc tggctcga ctcctgacct 1140
caagtgatcc acccgctcg gcctcccaga gtgctgggat tacagatgtg agccgccgcg 1200
cctagccagt tcactttctt aatgatgtct tttgatgatg gaaagtccta actgtaatgg 1260
agctcgcttt cccaatgctg actcttatgg ttagtgcctt tggagttaa gaagcatttc 1320
ctgctccaag atcatgaaga tactctcctc tgtcttatgg aagcttggtt atttttgcct 1380
tcacatttag atctttcatc taccacagat gaatgctacc tgcttttacc ctgagaactg 1440
tgtttggggg gaccatgtac ccctgagggg ctcttcgggg cacacagctc ttctcttacc 1500
atgggcctca gaggcaggcc cggagtgagt ttcagacttt gtgagtgaag cccttcaaaa 1560
cacgaaatat tcccagaaac ccagtaagtg cagcagacct actctaactg ggggcagtgg 1620
gaggacgccc acatcctgcc ccctcagccc ctctctgaca cccagggtg gccctgaatc 1680
caggggccct aggagcccag ctttagaatc accgcgctgg gtactcgatg gagcttgtct 1740
ctgatgcaga aactcctag cattctctct cagggtctt ttcatgtgaa tgacctagag 1800
gattgagctc atgtaggcac tgaaggcttc cacctctccc ataccgcaa ggccgatctg 1860
ccttcagctc ccagcaagtg tggggcagcg cgggccacag agtagggtgc agggatgggg 1920

cccctgcagc acccagggtc tctggtatgg agacagcagt gtggagtctg gaaactcaga 1980
gtcctttctgg ctgccgccgc ggctttacca tctggagagc caccacgctg aagcctcctc 2040
caccctgagc gcttggctgg cttcaggcct gtctcaagat gcaaggagag gatacaccac 2100
catcctgctg gctgctctga gtgtcacccc cctgaaagca gcacagggtg cccctcccat 2160
cctggcaccc cctacttctc ccccagtgga tgcagaatgt gctgggccag gtgctggacg 2220
cgctggaata cctgcacat ttggacatca tccacagacc cctttcgtaa gtcctggatg 2280
gccccctgaag ccctcaactt ctccttcagc cagaaatcag acatctggtc cctgggctgc 2340
atcattctgg acatgaccag ctgtctcttc atggatggca cagaagccat gcatctgcgg 2400
aagtcctcc gccagagccc aggcagcctg aaggccgtcc tgaagacaat ggaggagaag 2460
cagatcccgg atgtggaaac cttcaggaat cttctgccct tgatgtcca gatcgacccc 2520
tcggatcgaa taacgataaa gtctgccgtg gcccccgag ctggtggagg tgggtggtcac 2580
gaccatggag ctacatgaca gggtcctcga tgtccagctg tgtgcctgct ccctgctgct 2640
gcacctcctg ggccaagcgc tgggtcacca cccggaagcc aaggctccct gcaaccaagc 2700
catcacctcc accctgctga gtgtctctca gagccacccc gaggaggagc cacttcttgt 2760
catggtctac agcctgctag ccatcaccac aaccaggggt gtgtctgcca gccacctcct 2820
gccccaccca cgctccagga cagcccttcc caggggtctt ggaagggttg gtttggggta 2880
taggtgggtt ggacgggaca gtgtgggcc tcctcctgag atacatggtg gcatttggcc 2940
gtcttcattt ggccaccca aatgctggtc gcaccccttt ccatcttgat gacaagcttc 3000
cactcttgaa gtcactggtt ccctctacag acatgctagg cgcagctgtg ggcttcacac 3060
caatgacatc tctttccac acttctgcc ccttctggga ggctggggct caaatgcct 3120
gtgtgtctcc attccatagg gccagtggtt cttccgaagc cgccagccag gactgtggga 3180
aggagagggc catacagagc gctcacacct tcaccacaa atcgggtggg cactgttctc 3240
cccaacagga agctgggcct cgagagagcc taaggacagt tgccaggagt ccatgcagca 3300
gggttcaggg ctggggtctg ggccccagca ccctctttac tgcacagact ggataactga 3360
tgatacatgg ctgatctcac tttggggagt gaaaggaggc actaggaata gatgtcaact 3420
ggaaccctca ggcaaatgg atgtcagttc atcctaccgg gatagggcc cgtcatggtt 3480
ccatcctgga aggcacaggc tggctctgtg agccagggag gcagggtcag gccccctgga 3540
tggaagcta cagaggtcag acccagcctg gtagtgggat ggcagctatt gggactggtg 3600
gtccacgaga tggacagact cctctggggc cagtcccaca tcctcctgtt cagggtcca 3660

ttgagtgcac acgacttggc ccagagcagg cacctaggat tgcaggtaa atgggactgc 3720
agtgtccaag gacaacagag gcaggaaggc ttcttggagg gaggggccct ggggccctca 3780
ttcttggtca cccacagagt cagagtcact gtcagaggag ctgcagaacg ctgggctgct 3840
ggagcacatc ctggagcacc tcaacagctc cctcgaaagc agggacgtct gcgccagcgg 3900
cctgggcctg ctctgggccc tcctgctgga cgggtgagggg ccctcctcct gctgtccac 3960
cggggctggc agccctcccc cagccctcc ctaactgccc ctgagagcct tcgaggacct 4020
ccatgtcctg tccctaaaac acaacagcca tagtccggga aaggctcttc tgagagcttc 4080
caactccaac agaagaaaat cgaggagcag agagagaaaa ggcaggggag aaaggccttc 4140
tggcagaggc cgggtttcag gacttcttgc ccagtgggca gaccctcag ttttaagtgc 4200
ctcctgcca gggaaatgtc ctgggatttt ccgggcagtc ctggttccag agggcagcgg 4260
tggtgtggtg ctggatgcc tggtttgttt gatatttgat tcacagtagg gggccccctg 4320
gcctgtgctg cttcctctcc tctagacccc atcttggcac tccagcgccc caggaaaaag 4380
agagctccaa accacggaaa gcccgggaaa cccaagaacc ctgccagcac ccaaagtgtg 4440
ggattctcca agcctctcct gggctaacct ctgcacccgt ctctgaggac agttgacctt 4500
tcccaccca tttctgctgt tgctgtagc tggaggaagg cagcagatgg gggatgggaa 4560
ggccccctg cacacacca aggcctgggt gtccccctcc atccctgtcc tcgttccagg 4620
tatcattgtg aacaaggccc ctttgagaa ggtcccggac ctcatcagcc aggtgttggc 4680
cacctaccct gcggatgggg aaatggcaga agccagctgc ggagtcttct ggctgctgtc 4740
cctgctgggc tgcataagg agcagcagtt tgaacaagtg gtggcgctgc tcctgcaaag 4800
catccggctg tgccaggaca gagccctgct ggtgaacaat gcctaccggg gactggccag 4860
cctggtgaag gtgtcagagc tggcggcctt caaggtggtg gtgcaggagg agggcggcag 4920
tggcctcagc ctcatcaagg agacctacca gctccacagg gacgaccgg aggtggtgga 4980
gaacgtgggc atgctgctgg tccacctggc ttctatgag gagatcctgc cggagctggt 5040
gtccagtagt atgaaggccc tgctccagga gatcaaggag cgcttcacct ccagcctggt 5100
gagtacagc agcgccttca gcaaaccagg cctccctcca ggtggaagcc ccagctggg 5160
gtgcaccacg tctgggggac tggaatagat gtttgtatgg aactgacctt gatctccacg 5220
tgtatagttt tcaagactgc tctcctgect gcctattatc ccatctctat gactgggcca 5280
aaatcaatct taaacgggag gggtaatcag acctctccaa agagtttctt gtccatgact 5340
gctggattga gtcacatgag taactgctcc tggaccggg gactgtccac gaaaactgac 5400

ttgcctgctt cctccttcca ggaactgggt tcttgccgga aaaaagtgt cttgaggctg 5460
 gaggcagcca cctctcccag cccactgggt ggggaagcag ctcagccctg atgcggggga 5520
 gaagacagat accccacagg cccctccctc cacgtgtgcc ctctccctgt ccttcctttc 5580
 catgggccac tgtttccctt ggggtggggg gaagggtcat ccagcaccag aatgcgcacc 5640
 tcacactcct cttaggtgac taataaagag gcccaaggcc 5680

<210> 7

<211> 3590

<212> DNA

<213> Homo sapiens

<400> 7

aatcccagcg ctttgctgga ggatcgcttg agcccaggaa ttcaagacca gcctgggcaa 60
 catggcgaga ccctgtctgt cccaaaacaa aaaaacagat ttttaattggt ttgggatgag 120
 gcctggacac agggaaattt taaagctctc taggtaactc caatgtgcgt cgaggttgag 180
 aaccgccggc ctgtcacaca gagggagctc aggagatgct gagagtggga ggatgccggc 240
 ttgggagcct ggagttgcag cttcctgagg ctggcgagcc cctccttgct cccctctacg 300
 cctctgcatc gcacccacc cctgtacccc accttcctcc caccaaggaa acctcacctg 360
 ccgcctcccg cccaggtcct ttggattttg ccggtgtgtg tgggtgcgca caagaggccc 420
 ctccctgcca ggagagcagg tgagcctggc cgcgcacgaa tccgaggggg cgccgcacca 480
 gctgggaagc agccctgaaa tagacccccg ccgccccgc tgctccttc cggaatctgc 540
 tcagataccc cagatcagct acgcctccac agcgccagac ctgagtgaca acagccgcta 600
 cgacttcttc tcccgcgtgg tgccctcgga cacgtaccag gcccaggcca tgggtggacat 660
 cgtccgtgcc ctcaagtgga actatgtgtc cacagtggcc tcggagggca gctatggtga 720
 gagcgggtgt gaggccttca tccagaagtc ccgtgaggac gggggcgtgt gcatcgccca 780
 gtcggtgaag ataccacggg agcccaaggc aggcgagttc gacaagatca tccgccgctt 840
 cctggagact tcgaacgcca gggcagtcac catctttgcc aacgaggatg acatcaggcg 900
 tgtgctggag gcagcacgaa gggccaacca gacaggccat ttcttctgga tgggctctga 960

cagctggggc tccaagattg cacctgtgct gcacctggag gaggtggctg aggggtgctgt 1020
cacgatccctc cccaagagga tgtccgtacg agaccgtgag cgaattgggc aggattcagc 1080
ttatgagcag gaggggaagg tgcagtttgt gatcgatgcc gtgtacgcca tgggccacgc 1140
gctgcacgcc atgcaccgtg acctgtgtcc cggccgcgtg gggctctgcc cgcgcatgga 1200
ccctgtagat ggcaccacgc tgcttaagta catccgaaac gtcaacttct caggcatcgc 1260
agggaaccct gtgaccttca atgagaatgg agatgcgctt gggcgctatg acatctacca 1320
ataccagctg cgcaacgatt ctgccgagta caaggtcatt ggctcctgga ctgaccacct 1380
gcaccttaga atagagcgga tgcactggcc ggggagcggg cagcagctgc cccgctccat 1440
ctgcagcctg ccctgccaac cgggtgagcg gaagaagaca gtgaagggca tgccttgctg 1500
ctggcactgc gagccttgca cagggtacca gtaccagggtg gaccgctaca cctgtaagac 1560
gtgtccctat gacatgcggc ccacagagaa ccgcacgggc tgccggccca tccccatcat 1620
caagcttgag tggggctcgc cctgggccgt gctgcccctc ttcctggccg tgggtgggcat 1680
cgctgccacg ttgttcgtgg tgatcacctt tgtgcgtac aacgacacgc ccatcgtcaa 1740
ggcctcgggc cgtgaactga gctacgtgct gctggcaggc atcttcctgt gctatgccac 1800
caccttcctc atgatcgtg agcccgacct tggcacctgc tcgctgcgcc gaatcttct 1860
gggactaggg atgagcatca gctatgcagc cctgctcacc aagaccaacc gcatctaccg 1920
catcttcgag cagggcaagc gctcggctcag tgccccacgc ttcatcagcc ccgcctcaca 1980
gctggccatc accttcagcc tcattctcgt gcagctgctg ggcatctgtg tgtggtttgt 2040
ggtggacccc tcccactcgg tgggtggactt ccaggaccag cggacactcg acccccgtt 2100
cgccaggggt gtgctcaagt gtgacatctc ggacctgtcg ctcatctgcc tgctgggcta 2160
cagcatgctg ctcatggtca cgtgcaccgt gtatgccatc aagacacgcg gcgtgcccga 2220
gaccttcaat gaggccaagc ccattggctt caccatgtac accacttgca tcgtctggct 2280
ggccttcac cccatcttct ttggcacctc gcagtcggcc gacaagctgt acatccagac 2340
gacgacgctg acggtctcgg tgagtctgag cgcctcgggtg tccctgggaa tgctctacat 2400
gccccaaagtc tacatcatcc tcttccaccc ggagcagaac gtgccaagc gcaagcgcag 2460
cctcaaagcc gtcgttacgg cggccacat gtccaacaag ttcacgcaga agggcaactt 2520
ccggcccaac ggagaggcca agtctgagct ctgcgagaac cttgaggccc cagcgtggc 2580
caccaaacag acttacgtca cttacaccaa ccatgcaatc tagcgagtcc atggagctga 2640
gcagcaggag gaggagccgt gaccctgtgg aagggtgcgtc gggccagggc cacaccaag 2700

ggcccagctg tcttgcctgc ccgtgggcac ccacggacgt ggcttggtgc tgaggatagc 2760
 agagcccccga gccatcactg ctggcagcct gggcaaaccg ggtgagcaac aggaggacga 2820
 gggggccgggg cggtgccagg ctaccacaag aacctgcgtc ttggaccatt gccctcccg 2880
 gccccaaacc acaggggctc aggtcgtgtg ggccccagtg ctagatctct ccctcccttc 2940
 gtctctgtct gtgctgttgg cgaccctct gtctgtctcc agccctgtct ttctgttctc 3000
 ttatctcttt gtttcacett ttcctctctt ggcgccccg gctgcttgta ctcttggcct 3060
 tttctgtgtc tcctttctgg ctcttgcctc cgctctctc tctcatctc tttgtcctca 3120
 gctcctcctg ctttcttggg tcccaccagt gtcacttttc tgccgttttc tttcctgttc 3180
 tcctctgctt cattctcgtc cagccattgc tcccctctcc ctgccaccct tccccagttc 3240
 accaaacctt acatgttgca aaagagaaaa aaggaaaaaa aatcaaaaca caaaaaagcc 3300
 aaaacgaaaa caaatctcga gtgtgttgcc aagtgtgtcg tcctcctggt ggcctctgtg 3360
 tgtgtccctg tggcccgag cctgcccgcg tgccccgccc atctgccgtg tgtcttgccc 3420
 gcctgccccg cccgtctgcc gtctgtcttg cccgcctgcc cgctgcccc tcctgccgac 3480
 cacacggagt tcagtgcctg ggtgttttgt gatggttatt gacgacaatg tgtagcgcat 3540
 gattgttttt ataccaagaa catttctaataaaaaataaac acatggtttt 3590

<210> 8

<211> 4135

<212> DNA

<213> Homo sapiens

<400> 8

agaaaggcag gaagagaagc agtgaaggca gagctcactg gcccagctca gcagggtta 60
 agagaggcct ctcccctggt gatgtgtgtc tgaggatgag ggccccctgc tttcttgggg 120
 gcagtgtcta ccaggatcct ctgctctgga actcagaaag gaaagagggc tcttgagta 180
 gaaggcaagc tcctgggaag tcaaggcagc cagagatagc tctgggtgcc ggcacctcg 240
 aagcagccac ccttcatgga gaaggggaca gagctgctgg tgtccccatc ccagagtgga 300
 cctgggggagc atcagcccct cctcgtgaag cacaggagg acagcgccgt ggtactgagt 360

ctcatcgaca ctattgacac cgtcatgggc catgtatcct ccaacctgca cggcagcacg 420
ccccagggtca ccgtggaggg ctctctgcc atggcagagt tttccgtggc caaaatcctg 480
cccaagaccg tgaattctc ccattaccgc ttcccggccc acgggcagag cttcatccag 540
atccccacg aggccttcca caggcacgcc tggagcaccc tcgtgggtct gctgtaccac 600
agcatgcact actacctgaa caacatctgg ccgcccaca ccaagatcgc ggaggccatg 660
catcaccagg actgcctgct gttegccacc agccacctga tttccctgga ggtgtcccca 720
ccaccacccc tgtctcagaa cctgtcgggc tctccactca ttacggtcca cctcaagcac 780
agattgacac gtaagcagca cagtgaggcc accaacagca gcaaccgagt cttcgtgtac 840
tgcgccttcc tggacttcag ctccggagaa ggggtctggt cgaaccacgg ctgtgcgctc 900
acgagaggaa acctcaccta ctccgtctgc cgctgcactc acctaccaa ctttgccatc 960
ctcatgcagg tgggtccgct ggagcttgca cgcggacacc aggtggcgct gtcgtctatc 1020
agctatgtgg gctgtccct ctccgtgctc tgcctggtgg ccacgctggt caccttcgcc 1080
gtgctgtcct ccgtgagcac catccggaac cagcgctacc acatccacgc caacctgtcc 1140
ttcgccgtgc tgggtggcca ggtcctgctg ctcatagtt tccgcctcga gccgggcacg 1200
acccccctgcc aagtgatggc cgtgctccta cactacttct tcctgagtgc cttcgcatgg 1260
atgctggtgg aggggctgca cctctacagc atggtgatca aggtctttgg gtcggaggac 1320
agcaagcacc gttactacta tgggatggga tggggttttc ctcttctgat ctgcatcatt 1380
tactgtcat ttgcatgga cagttacgga acaagcaaca attgctggct gtcgttggcg 1440
agtggcgcca tctgggcctt tgtagccct gccctgtttg tcatcgtggt caacattggc 1500
atcctcatcg ctgtgaccag agtcatctca cagatcagcg ccgacaacta caagatccat 1560
ggagacccca gtgccttcaa gttgacggcc aaggcagtgg ccgtgctgct gcccatcctg 1620
ggtacctcgt gggctctttg cgtgcttgct gtcaacggtt gtgctgtggt tttccagtac 1680
atgtttgcca cgtcaactc cctgcaggga ctgttcatat tcctctttca ttgtctcctg 1740
aattcagagg tgagagccgc cttcaagcac aaaaccaagg tctggtcgct cacgagcagc 1800
tccgcccgca cctccaacgc gaagcccttc cactcggacc tcatgaatgg gacccggcca 1860
ggcatggcct ccaccaagct cagcccttgg gacaagagca gccactctgc ccaccgcgtc 1920
gacctgtcag ccgtgtgagc cgggaggctg ccaaccaggc caggctgcgc tcagaacaca 1980
ccccccaaa cagaatgaaa tgccccacct ttgcccattg accctctcct tgctgctgtc 2040
tggacatggg tgttgtggcc ccgagacagc tgtcctcccc tgtgactctg gctgtcggag 2100

cacactgctc agcccagcag cctgatgccc aggccagcgt gggccctcct gccttgcac 2160
caccctgtggg ctgagtact tcctcggggg attcccagga cacagtggcc tgactgtgat 2220
ggtgcccttg agcctccctt catcactcag catcagaccc agcgaggcca ggacactcgg 2280
ggccggtccc gcagcaccag gaggggatgt tcagcctctg tgccttgggtg gggcttgggg 2340
actcagggcc aaagaggtgg ttcaggtccc cacgcaccct cagtcaggcg caggcagctg 2400
gggggtgtgtg gggaagagca tgcggagtcc ccagtgtctg aatccactga gtggtgagtt 2460
ccccacagcc ggcgctagcc gtggtgtgtg tctctgtagg tgggtgccggc gtgggccaac 2520
ctgtgtgtgtg tcatcagttg ggggcccctg cccaagccga gctcgagccg tgggcgggag 2580
tcgttgactc tccaggtgag ggcgaccctt ctgccctgtc cttggggggg tcccctctgc 2640
tcacgtgaag agccgctctg ggccttgagg ctgcctgatg gtgcctgtgc ttggggggagc 2700
ttctcggcca tccgctgtga gttttgcctc tttggacccc aattcggcct taagatgccc 2760
tcctccctcg tgtgccagcc tccttggttg ttcttgggcc acaggagctg gccgtgtccc 2820
cgcagtgcct ggtgtccagg tggaaattgg agggcatitt ccagggcact gctttcccca 2880
gaggcttcct catggctcac aggcactcta cgaagtttct aatgggcaga ccacgcggca 2940
ggtagcacag tgcgtccgt ctggtcacca tgagaccgac ctgcgtgag tccccactga 3000
cctggagagg gagggctggt gacagccgtg tcttctgtgt tgagggaaat ttatggactc 3060
agattcagcc ccagaggaga tgggataatt gttatggacc catgtgtggg catgatcctg 3120
tggaacacag gtttgggatc atagatgtga attaagacac caccgagata cgggctgtga 3180
ggttcatact gtgctgatag cactcgtggt gtctgtgaaa tgtgggtaag acattcaaac 3240
ctggttttga tactggaac tcttccttta aaactgtgac catgatttca ttcagcccct 3300
ccacaccctt atgtctgcct tgtttcagag tgagttttct atggagcctg tggccctttt 3360
gcagcccacc tgggtggcttc ttaatgtaac tcttcccctg gtcgcctgga gtggaccact 3420
catctgcagg cctctcctgc atggggaggg taggcaggga gcagcatgtc tgcaggggtg 3480
aacctttgct cttctgtcag gcgaggccca ggctgcacca gccacctgcc acatggtgac 3540
agtgccacgg gccctgcgta tggcccctgc aaccgtgtc tggcgggcac acctggctgc 3600
tgcaggccaa ggccgctgtt cagtgaagag tcccatgttt agtatggact aaagtcccat 3660
gttttagccac tgccccaggc tcccgtagcc ccagaaacca ggtcacatgg accacagtgc 3720
cagatcctca tcacgccggt gagcacctag aagtgagaac actgtattcc tacaatgtac 3780
acttggatat ttctccttat ttagtttcta gtgaaacaaa tcaagtaagg aactatcttt 3840

agtttagatg gaattatttg tttttaattg ttgccgtatt catctatata gctaataattt 3900
caagataagt aatgaacaaa acctgtctaa accttttggt tccaatgaat gaaagtcag 3960
cactttatatt ataggctcta tgttttggct tctgcagtac ttttattatc tatacataat 4020
ttggccaaaa ataagaaatt ggaaagaatg aaatgtttag tttatagtag aagaaagatg 4080
atgacactaa gttgtgaaaa tatgttgtga tttttatgaa ataaactcat gtcct 4135

<210> 9

<211> 3811

<212> DNA

<213> Homo sapiens

<400> 9

tataaacagg aaaatttaaa tattgtagtt agagtttttt tcttcttctt agagatgggg 60
tttcaactctg tctcctaagg ctggagtgcg gtgggtgggat catagctcac ttcagcctca 120
acctcctggg ctcaagccgt cttcctgcct cagcctcccg agtagcacgc atcaccagcc 180
cagctaattt tctgtctttt tttttttttt tttttagtag acgggggtctc cctgtgatgc 240
ccaggttgtt cttaaacttc tggcctcaag tcatcctcct gtctcagcct cccaaagtgt 300
tgggattaca ggcatgagcc accgtgcctg cctgcagtta aacattcccc caaaaatac 360
tgggccacac cgaggctagt atgaaagttg cagccctggg tatgtgtgac ctgagagccc 420
cttggcagcc atggggagcg atgggggaag ctggggcact gggaaggggt ctaggcacac 480
gagccctcgg gtcctcacct cccacgaggc aggggtcccag gctgcctcta ctcaaggtgg 540
ggcagcaaaa gaatcgccct ctagctgtgc agaccccctc cccctccagg tctgcacacc 600
gccaaggct ggccctggac ttctggtgga ttcaagtact caggtgggga gagtaggagg 660
ccagtcgcac ggagccctcc ccaacgcaca gtgcctgccc agctcaggtg cctctgaggt 720
gtagattctg agcaggttct gggatctcct ggcccccagc gggggctgcc acaggagtat 780
aggcccaaca ggaagggaag atcccagcga ttggttccag tgggttctgg ggatgtcctc 840
ggagcctggc ctggctcagg gtccctccaa aagtccctgg aacgagggtg aagtgttgct 900
catgaatcag agagaaagca agcaggaccc tcaagaaggg aggagagaga ccattgggcg 960

tgcagcactt tgcactcggg ggaagggctc ctgaaacttg gaagtcagca aatatctgct 1020
gtacttcacc agaggtggga tctaaggaga cagaatgaag ccagggggca gaatgtttcg 1080
tttctgtctg aaacctgtgc cctgcatccc tcatggactg gggacctgga gttcccgggg 1140
accaccgtgg ggaaatgcgt cccagcacta atggccccca gcagctgcag tcacgggggtt 1200
tgggcgctct cccctcctga tttgggggac cttttgtgct cctctgggca gagggaggag 1260
gcagagggag gaggaaggcc cttcgctgtg ggctgagtcc ttcccacctt ccataccagc 1320
ccagcaggaa gccactgcag gatgccccag aggacagcct gatggttggg ggagaggctt 1380
ctcccgccct caccctccg ggttcttctt ggaccactg agtaaccag gtggtgggac 1440
gtgtggctgt gagtcttggg ctctgggcgc tcaccagcag ctccgtctca cacctgcctg 1500
cgtgtgacac cagcaccct tatgatgagg aaacaagttt gccagctgc aggggtggcc 1560
gagtcaggat gactctagcc cgtgtcctgg acaccagacc ctgcccaggt ccagccgggg 1620
ctggtctcag ctttcttggg ctatgtcgcg gaggggtgtt gggacagcga gaggctggcg 1680
tggacagtgg aggggtgact ttgtgggtgg tcctgatagt gacggagagg aggatactca 1740
gctccacccc tgggcggccc ctgagcagca cggctgggccc tgaaggggca gggctgccgt 1800
cacaggctct ggcccctggg agctcaaggg gtgaatccct gatcccaggt gtgggactgg 1860
gatggggcct caggctgatg caggcaggac ctccagagct caggactggg tgggtgggct 1920
cacagggagg taggggcagg ccagagtccc agctgtcctg gactctgctg tggggaaggg 1980
ctgatgcagg tgtggagtca aatgtgggtg cctcctgcag ccgggtgcca ggaggggtgg 2040
aggggccacc ctgggctttg tccgggagcc tggctcttccc gtccttgggc tgacaggtgc 2100
tgctgcctct gagccctccc tgctaagagc tgtgtgctgg gtaaggctgg tggccctttg 2160
ggctccctgt ccaggatttg cgctctggag ggtagggctt gctgggctgg ggactggagg 2220
ggaacgtgga gctccttctg cctcctttcc tgcccatga cagcaggcag atcccaggag 2280
agaagagctc aggagatggg aagaggatct gtccaggggt tagacctcaa gggtgacttg 2340
gagttcttta cggcacccat gctttctttg aggagtittg tgtttgtggg tgtggggtcg 2400
gggctcacct cctcccat cctgcccag aggtgggcag agtgggggca gtgccttgct 2460
ccccctgctc gctctctgct gacctcggc tccctgtgct gcccaggac catgaatggc 2520
acctacaaca cctgtggtc cagcgacctc acctggcccc cagcgatcaa gctgggcttc 2580
tacgcctact tgggcgtcct gctggtgcta ggccctgctgc tcaacagcct ggcgctctgg 2640
gtgttctgct gccgcatgca gcagtggacg gagaccgca tctacatgac caacctggcg 2700

gtggccgacc tctgcctgct gtgcaccttg cccttcgtgc tgcactccct gcgagacacc 2760
tcagacacgc cgctgtgcca gctctcccag ggcattctacc tgaccaacag gtacatgagc 2820
atcagcctgg tcacggccat cgccgtggac cgctatgtgg ccgtgcggca cccgctgcgt 2880
gcccgcgggc tgcggtcccc caggcaggct gcggccgtgt gcgcggtcct ctgggtgctg 2940
gtcatcggct ccctggtggc tcgctggctc ctggggattc aggagggcgg tttctgcttc 3000
aggagcaccg ggcacaattt caactccatg gcgttcccgc tgctgggatt ctacctgcc 3060
ctggccgtgg tggctctctg ctccctgaag gtggtgactg ccctggccca gaggccacc 3120
accgacgtgg ggcaggcaga ggccaccgc aaggctgccc gcatggtctg ggccaacctc 3180
ctggtgttcg tggctctgctt cctgcccctg cacgtggggc tgacagtgcg cttgacagt 3240
ggctggaacg cctgtgccct cctggagatg atccgtcgcg ccctgtacat aaccagcaag 3300
ctctcagatg ccaactgctg cctggacgcc atctgctact actacatggc caaggagttc 3360
caggaggcgt ctgcactggc cgtggctccc cgtgctaagg ccacaaaag ccaggactct 3420
ctgtgcgtga ccctgccta agaggcgtgc tgtgggcgct gtgggccagg tctcgggggc 3480
tccgggaggt gctgcctgcc aggggaagct ggaaccagta gcaaggagcc cgggatcagc 3540
cctgaactca ctgtgtattc tcttggagcc ttgggtgggc agggacggcc caggtacctg 3600
ctctcttggg aagagagagg gacagggaca agggcaagag gactgaggcc agagcaaggc 3660
caatgtcaga gacccccggg atggggcctc acatttgcca ccccagaac cagctcacct 3720
ggccagagtg ggttctctgct ggccagggtg cagccttgat gacacctgcc gctgcccctc 3780
ggggctggaa taaaactccc caccagagt c 3811

<210> 10

<211> 2615

<212> DNA

<213> Homo sapiens

<400> 10

agaaatgttt cgagctgggg cagagggagc agagaaggag ccctcaccgc ggccggaatg 60
cagagcggac cctggcccag gactgggttt ccctttaggc tcgggcctac cctggccctc 120

gctgttggaa tctccaggag gccgcatttt agatatcccc tgcaaagtgt gtggcgaccg 180
cagctcgggg aagcactacg gggctctacgc ctgcgacggc tgctcagggt ttttcaaacg 240
gagcatccga aggaatagga cctatgtctg caaatctgga aaccagggag gctgtccggt 300
ggacaagacg cacagaaacc agtgcagggc gtgtcggctg aagaagtgtt tggaagtcaa 360
catgaacaaa gacgccgtgc agcacgagcg ggggcctcgg acgtccacca tccgcaagca 420
agtggccctc tacttccgtg gacacaagga ggagaacggg gccgccgcgc actttccctc 480
ggcggcgctc cctgcgccgg ctttcttcac cgcggtcacg cagctggagc cgcacggcct 540
ggagctggcc gcggtgtcca ccactccaga gcggcagacc ctcgtgagcc tggctcagcc 600
cacgccaag taccatcatg aagtgaatgg gacccaatg tatctctatg aagtggccac 660
ggagtcgggt tgtgaatcag ctgccagact tctcttcacg agcatcaagt gggctaagag 720
tgtgccagcc ttctccacgc tgtctttgca agaccagctg atgcttttgg aagatgcttg 780
gagagaactg tttgttctag gaatagcaca atgggccatt ccggttgatg ctaacactct 840
actggctgta tctggcatga acggtgacaa cacagattcc cagaagctga acaagatcat 900
atctgaaata caggctttac aagaggtggg ggctcgattt agacaactcc ggttagatgc 960
tactgaattt gcctgtctaa aatgcacgt cactttcaaa gccgttccta cacatagtgg 1020
ttctgaactg agaagtttcc ggaatgtgc cgccattgca gcccttcaag atgaggctca 1080
gctaacgctc aacagctaca tccataaccag atatccact caaccctgtc gctttggaaa 1140
actcctgttg cttttgccag ctttacgttc tattagccca tcaactatag aagaagtgtt 1200
tttcaaaaaa accatcgga atgtgccaat tacaagactg ctttcagata tgtacaaatc 1260
cagtgatatc taagtcaca agataccac ttttcaggat gggacagtat cagatgaact 1320
tcaacctatg gagaacaagc ctcaactaac aaacccttca ggaagcatat accggggaat 1380
gtgtagcctt caggaaaaaa atgccaattg acacaaagca ttccagtagc tatgacctgc 1440
cgccctgacc aggatagggc ggggtgggaag gagaggggtg caacaggacc gcctgcactg 1500
aaaactcact gctgccatgc cctgggaggg ggcaaaactgg gggttgccac aggccgtgcc 1560
attctgcctc ttacctggaa gatcaggctg aacgatcaaa agctgaaaca taagtagtgc 1620
tttctcttc tttttagcat acaaagtttg gtaaccaa atagctctgt gtataacatc 1680
gtactgcggc cttcaaaact acgttatgtt ggagcattta ttttaaaaat aatggtaggt 1740
tttaaattaa aagtgttatc aaaagtttcc cctctattgt aatacattat taagtggcct 1800
tcagaactga gttaataagt gaaaagtagc ttatgccatg tgatttgctt tttctctatc 1860

```

ttcttttttc ttttctttct ctttcttttg ttctttcttc ttctttttta ataccatagg 1920
ccaggcaacc ttgtcaaagg aattgatgga caaaatgaga ttctcaccag gacttcaggt 1980
tggaatacat gtcaaaaaga gaagagcttt actaaaagaa cataagtcaa gggaggatga 2040
ataaaaacag caaaaccaa taaagtggag atgagtgatt gagtgaggta gattgctgtc 2100
ccgttaacat agtgctgaaa ccaaaggcag tgggggtcca aactcgtggt gcagcaagtc 2160
acaccagaca ggaaacgaat atggacgtaa ttgcagaagg aacttcaagg agatgaatga 2220
atgctaaaag ggttcttgac tgatccattg ctaacagcct gagactcttc aatgcctttc 2280
cgaaaactgg tttctagtaa agccttgaat gaacacacac acacacacac acacacacac 2340
acacacacac acacacacac acacaccgtc ctacacttta agctgctcct ttggtatgac 2400
tctatactta tggaaatgta gaaacaaaca tttttaaata gctgctgtac ttttcacatt 2460
ttgatttatt aggtactagc actgaggaaa aaaattcagc acttggacta tcataggatg 2520
agtaaaactt ttcttgtaca aagtgaatta actgatttgt gaagttaaaa ggttgtactc 2580
attgtattta caaagaataa aaatatattg aattt 2615

```

<210> 11

<211> 3794

<212> DNA

<213> Homo sapiens

<400> 11

```

ttgggttttt aatacaggtc gagagctatg aattttgatt ttccttttaa aaagggatca 60
ggaatacctc tactaacaac caatttgtcc ccacaatgcc tctccctcct gcacgcaatg 120
gtggcctatg atcccgatga gagaatcgcc gcccaccagg ccctgcagca cccctacttc 180
caagaacaga ggtaggcact ctcgtgggtt cgggtggggca gatgcaggac ctggcaaatg 240
agacccatgg gcacatgggtg gggcaaggta ctgccccctg gagggtcctt ggttcttggg 300
aatcgcgctc tctatcctga agagtgggtc gggcctcttg gggccacac cgtgctgact 360
tctggtgtat cctcccaaaa atgaggatgt tagcagctcc acggcaggag aaagaggggc 420
tggtgtcca ctttgcggtt gatgccaagg ctacaggtag taatccttct agaagagcct 480

```

ggtgtgcggc ccttgtaaat ggtgtgaaca ggatgcattc tgagatggta gagcactgtg 540
gatgtcacct ggcgctggaa gtacctgctg aatgtgaatc taagacctgt ctgccctgag 600
gtcacctaata tccggacctc tggacactcc tgtttgccta actgtggagg tgacagtttg 660
atctgaaact gcgagtcctt gggtaggcag cgcttctgat tcatggagcc agctctgttc 720
ctggcccaga gactgtcagg ctgagacaga aaagaaaaaa aaaggctcct gattacagct 780
caggcagaaa agaagagcta gtgacaactc acacagtaga caaattggag acaaagaagc 840
cagtaggccg tgtgctctgc gggctctcag tggagctgct ccattctctt ctccttcccc 900
gtaggaaaac agagaagcgg gctctgggca gccacagaaa agctggcctt ccggagcacc 960
ctgtggcacc ggaaccactc agtaacagct gccagatttc caaggagggc agaaagcagg 1020
tactaagcga aatcggagcc ggagactgcc tagctcagcc gggactggat gccctcgggt 1080
ggcctgtgtc tctctcagga ggacgccacc tgtggaggct gtggcagacc acattgtcag 1140
tcgaggcctg aacagcatgc gagttgatgg ctccatgtgt tgggacatca gggaccccag 1200
ttctatittg tcaagccctt tctgccccag tagctaagaa ttgatgacga atatcccatg 1260
tgagcaggcc cagagagggg atgtttgcat gaacttttgt gtctgtcctg atgaccatat 1320
gcagggtcca ggttctatgc caaagcaccg tgcagacatc gcccctcaga acggccctgg 1380
agggaggcgc caagaagatg caatgctggc ccaagtcacc caccagcca gtgacaggtc 1440
tcaagtcact gtctctgcct ggctccagat gagcacttcg tcccatcagc agaggttgtg 1500
cacattttcg gttgagaaat gccagctttg cccctcagtg ccttcacctt cccacctacc 1560
gtgtggcatc ctacgttgt gtagccctgt gtcacctgct ggcagggtca caccagtctc 1620
ccagttctgg gctggcacc caacagtgtg acagtgtgcc ctccatcat tcgtctgttc 1680
ggtcgtccgt acgtcactcg ctctccag cctttaccga gtgtccctg accgcttggg 1740
agctgggaac ccagagctgc agtcaggctg ctctccagaa agccaggctc ctgtgtccca 1800
ttggtcctg atttgacagc tccacttgaa tgtccgaaag acagacttgc tgaaaccct 1860
tctcctgatc tttcccaaaa cgtgcttttt ccctgtggct tcccatggcc ttcctccga 1920
gggaggcaac tccgtcaact gccaggccc cccgtggacc ctctctcat cccacagggg 1980
cttcgccttc ttgctccta catctctctc accttgcta agtcaccaca gatggcctcc 2040
tactactgt cctgtctctg cctccgctcc ctctggaca gagggtctgc ttcaaaccct 2100
tgatgtcact cctggccggc attcacccaa gtccccacc tcagtcacat ccgaggctcg 2160
catatttttc tctaagggtc tgataggaaa tattttaggc cccgtggccg tggggtgttc 2220

gttacactcc ctgctttgct gcagtgtgaa aggagctgga ggcaacacat gaaccacata 2280
ggggctgcgt tccagtagaa ctggggacac taaaatttga atttcacata attttcacat 2340
gccatgaaac attttttttt cagtcattaa aaataaaaaac caggtgggtg cggtggctca 2400
cacctgtaaa cccagtgcct agggaggcca gggtaggatc acttgaggcc aggagtttga 2460
gaccagcctg ggcaacgtag gcaagaccac tgccccacca ccccatctgt tttaaaaatt 2520
agctgggtgt ggtgcgtgcc tttggtccca gctacttggg aggctgaggc aggaggatcg 2580
cctgagccta ggaagctgag gctacagtga gccaagattg caccactgca ctccagtctg 2640
tgctacagag caacaaacac cgtgtctcaa aacaaacaaa caaacaacaaa accccaaatg 2700
taaaaccatc cttagctcac aagccaccca aaaccaggcg acaggcagat cagggtgtga 2760
gaccttggag ctagagcagg aggcaggagg cgtgggcacc acagcccgcg agccctttgc 2820
cgcccgcact gctcagcttg cagtcggcat tcttcaggtt cccctgaag gtttggcctc 2880
ccccgtggga ctctagtggg cagccactcc cttgggggtgt ctgggacctg gtggatgtgt 2940
cccagtcaca tcttatggtc aagaaattgt attaggcaaa agcagtacaa aagagcctct 3000
ttcgtttttc ctaatttcca taaagcagaa agcgttttat gtccaactgg cctctgttct 3060
gtttcaagaa acagtcccta aagcaagagg aggaccgtcc caagagacga ggaccggcct 3120
atgtcatgga actgccccaa ctaaagcttt cgggagtggg cagactgtcg tcttactcca 3180
gccccacgct gcagtccgtg cttggatctg gaacaaatgg aagagtgccg gtgctgagac 3240
ccttgaagtg catccctgcg agcaagaaga cagatccgca gaaggacctt aagcctgccc 3300
cgcagcagtg tcgcctgccc accatagtgc ggaaaggcgg aagataactg agcagcaccg 3360
tcgtctcgac ttcggaggca acaccaagcc cgaccgggcc aggcctgggt gatctgctgc 3420
tgagacgcca cggagggtg gggatgcgcc tgcgtccgtt tcgcgctggc cggggctctg 3480
ggtgctgccc tgcgccctgc cgcacccgcg gcccgcgcag ctgcctagga tgttctgggc 3540
taatatactt gtaaaaccac cgcattctag ggttttcttt cattttcgtt aagaatttgg 3600
ggcaggaaat actttgtaac tttgtatatg aatcaaaaca aacgagcagg catttctgtg 3660
atgtgttggg cgtggttggg aggtgggttc tgcgtgtccc ttcccagcgc tgctggtcag 3720
tcgtggagcg ccatcatgtc ttaccagtga cgctgctgac acccctgact tttattaaag 3780
aataagctgt cggt 3794

<210> 12

<211> 4147

<212> DNA

<213> Homo sapiens

<400> 12

ctcggcgccc	ggggccggct	gcgggaaacg	ctcgcacaga	cagagacagg	cgcgcacccg	60
ggggcgccacc	cggggccaca	cacgcctccc	caggcgatgg	ggtccaggcg	gccccgcccc	120
aatgcttccg	tccggccagc	ccatccagcc	gctgccccgc	acctgtcccc	gccgctccgg	180
acccgcggcc	gccccccg	cacgcacacc	ggggacggga	cccggcgccg	cgactcccca	240
cgccccgcag	gcgacgccac	cttcggccgt	agcagcccgg	ggtccgaggc	atccccctc	300
ggagctggcg	cgggaccag	cgaacgggcc	ccggggcgga	gggacgcttc	gggacccac	360
agccgccgtg	ggtcgggcac	cgggggctcc	aggaggctcg	cggaggggtc	gccactcacc	420
agctgaaaag	cgcagaggac	gacgagcgcg	cagcggccgg	agcaggagcc	catggtgccc	480
gcctatacag	gaggcccccg	ggtgccccgc	gctcggcccc	cgcccccgct	ccgcccgtc	540
cccacgcgcc	gcagcctcga	ccccgcccc	tccgcatccc	gttccccct	cctccccga	600
acccccggcc	cagccccagc	cccgggtaac	agctgcgctg	cgtctcccc	cccaccccc	660
cctgcggccc	ctgggggctt	cgcgccggcc	ggagaggggg	acccttctta	aaggggcgat	720
ggcgcgccct	ctcgggtagc	gaggggcggc	gccgcggccc	caggacagac	ctgcagctga	780
ggggagcccc	ccccggacgc	ggcagggatg	accggcgctc	gccccctcct	gcaggatgac	840
ctgcccagca	cctgctcccc	ccgtgcgccc	tctggcgatga	gatccccatc	aaccttcccc	900
gtgtggggtc	tgagcgaccc	gcatccgcga	ggcgttttct	ctgagagtct	cagagattcc	960
ttggcttgga	gaagtcctgg	acccgggtcc	agagctcaga	ggtccccctc	acgggcgggc	1020
gggccgagga	ctcaggagga	gggagtgagc	cgggctcgcc	cacctccac	agggagctcc	1080
tctgctggaa	ggctgagacc	tgaaggccag	gtctgtctca	actgcctgga	ccggcctcct	1140
gcctgcgcgg	gaggggccta	gggagcctcg	ggttgggaca	gcctgaggcc	ttgtgcacgg	1200
tggaggcagg	accagctgc	ccagggcagg	tcctagggtc	tggatggggc	tgaaccccg	1260
aaagcctgcg	cttccggaag	aagcttgccg	tggaggcagg	gaagggtccc	agccaggccc	1320
gagaggagca	ggacgtggat	acagtgcgac	agagtgtcag	ccggtcggaa	agcccatggg	1380

ggcagaacag cgggagatca gggcaggtgg ccgggccagg gctgccactt ggtctttact 1440
cagtggcagg gcaggtcccc tgactgccac ccaccgcctg aggagcctga gatgggtttc 1500
ccatggcaag ctgaattcct gcccactcg agattttatg aattactgag gattcttgcc 1560
ttcctaattt tggttgtgtg tcccacagtc ccggggatca ggaggggtta acagaagacc 1620
aagggtcagg ttttctgcac tgcggggagc caccacagat ctgtagtcca ggctcatggg 1680
cgggtgaggg ggtagcttcc aagccccaaa ggtagagag caaaggtgtg gccagtgtgtg 1740
gagggagggg gagcgtcggg ttgtttgcag tgctggggag ataagaatat tcaggaatca 1800
gatggtggtg atggttgac agctctgtga atatactgca gccactgag cactacactg 1860
taaaggggtg actattatgg tgtgaggatt atatccgat ttttaaaga tgtgaaaaa 1920
ttaatgttta ttatacatg atataagaca ttgattttta aatacagcgg taccatcccc 1980
gtccccata accctgggtc ctccaggcac ttggagccat tgggtgaagg acgtggactc 2040
ctagaaggtg attcgcaaag gtgccacgta agcagaggag ccgggtaagc agaggcccag 2100
tgccacgtaa gcagaggaag gcccgagacc ggctgctgag gtcctgaggg agagggaggg 2160
gcacctgctc agccttgctt gtgtgtgtgg atccttctct ctcaggggaa gctgcctgga 2220
catgcatgcg ccttgagtt ctgtctgcc aggtgcctg ctcttctct tgaaagcaaa 2280
agacttcact tgctgaggat gcaacgatgc tctagtgtcc cttgcaggag ccagtgtcag 2340
agccaggagc cgtcgtgggc atccccgtc ccaaagtac gagttctggt gaaggggtca 2400
ggacagcccc tttagagatt cagcagagca ggtctgcaca ggtccacac cgtgcacagc 2460
ccacactcca cctgccctgg tgcacactga cggcagcagc cagccgcggg gaccccaggc 2520
gcagtgttgc ctcgtgaacg ttagagcagc ctccagaggc gggtttggtc tgcacacctg 2580
gttcgcgggt gacacagttg tgacttgcc caggtcagtt gttggggcct ggccagctgt 2640
gggtggtgag gttttctcgg acgcttcac tgctctcgtt cttgttcgaa ttccttctgc 2700
ctaagaatga cagctttaaa agtgaaatgg cagcacgtct ccatcattac acagaatccc 2760
atgtgggaag accctagtcg gagggcccca cccttgcccg gaactggact ttggggcagc 2820
agcaaactgc agtgagtgtg agtcacacct gaggacaggt gggctgtctt caggctccca 2880
acaggtctgt cgcttaaggg gctggggaat gggggctgga atgccacggc ctgcccacct 2940
gggggactct gggcagagct ggagcagcca ggcaggaaaa tatccccag tttccagaga 3000
gtaaaagctg atggtcctgc tgaggtgtga aggacagtga cggacagtga tggaaggtga 3060
tggaacagagg aggacagtga tggacagagg agggcagtga tggacagtga cgggcagtga 3120

tagacagcaa cagagagtgg aggacagtga tgaacagcaa cagacagtgg agggcagtga 3180
 tggacagtgg aggacagtgg aggatagcag agggcagtga tggacagtaa tggacagcaa 3240
 aggacagtga gggacagcag aggacagtga tggacagcgg aggacagtga tgggcagcgg 3300
 aggacagtga tggacagtgg aggacagtga tggagagcag aggacagtga tggacaatgg 3360
 aggacagtgg aggacagtga tggacagtgg aggacagtga tggacagcgg aggacagtga 3420
 tggcgagcgg aggacagtga tggagagcgg aggacagtga tggacagcgg aggacagtga 3480
 tggcgagcgg aggacagtga tggacagcgg aggacagtga tggacagcgg aggacagtga 3540
 tggagaacgg aggacagtga tggacagcgg aggacagtga tggacagagg aggacagtga 3600
 tggagagcgg aggacagtga tggacagcgg aggacagtga tggacagtgg aggacagtga 3660
 tggagaacgg aggacagcga tggacagcgg aggacagcga tggacagcgg atgacagtgg 3720
 aggacagcgg atgacagtgg aggacagtgg aggacagcga tggacagtgg aggacagtga 3780
 tggacagcgg aggacagtga tggacagtaa tggacagtga tggacagcgg aggacagtga 3840
 tggacagcgg aggacagcgg aggacagcga tggacagtgg aggacagtga tggacagagg 3900
 aggacagtga aagacagaga aggacagcaa tgggaaggaca gccaccccag tcagcacaaa 3960
 cgcaggccga gtggggccca tgcattcctg ggaggactgg caagctggga aaggcctcca 4020
 gggcaggga actcaggaga gtgacagctt gttcaggcca ggaccttgtg catcttctcc 4080
 ttttttcttt cttagtaaac ccagtgtatt catttttagc ggctgccata acaaattccc 4140
 ataaact 4147

<210> 13

<211> 3121

<212> DNA

<213> Homo sapiens

<400> 13

gtcgctactg atcatgcttg ggccagggtc caatcgcagg cgccccacgc aggggggagcg 60
 agggccaggg tccccggag agcccatgga gaagtaccag gttttgtacc agctgaatcc 120
 tggggccttg ggggtgaacc tggtggtgga ggaaatggaa accaaagtca agcacgtgat 180

aaagcagggtg gaatgcatgg atgaccatta cgccagtcag gccctggagg agctgatgcc 240
actgctgaag ctgcggcacg cccacatctc tgtgtaccag gagctgttca tcacgtggaa 300
tggggagatc tcttctctgt acctctgcct ggtgatggag ttcaatgagc tcagcttcca 360
ggaggtcatt gaggataaga ggaaggcaaa gaaaatcatt gactctgaga tcaagacaga 420
aggcgttcct aacactagaa ggttcccagt cggtgaccat gattccagat tgttctgccg 480
gtccctgaag ttcctagaca aggttttgct atgttgcccg ggctggcttc gaactcctga 540
cctcaagtga tccaccgcc tggcctccc agagtgtgg gattgcagat gtgagccgcc 600
gcgccatgcc agttcacttt cttaatgatg tcttttgatg atggaaagtc ctaactgtaa 660
tggagttcgc tttcccaatg ctgactctta tggttagtgc ctttggagtt taagaagcat 720
ttcctgctcc aagatcatga agatactctc ctctgtctta tggaaagctt gttatTTTTg 780
ccttcacatt tagatctttc atctacccca gatgaatgct acctgctttt acctgagaa 840
ctgtgtttgg ggggaccatg taccctgag gggctcttcg gggcacacag ctcttctctt 900
accatgggcc tcagaggcag gcccgagtg agtttcagac tttgtgagtg aagcccttca 960
aaacacgaaa tattcccaga aaccagtaa gtgcagcaga cctactctaa ctgggggcag 1020
tgggaggacg cccacatcct gccccctcag cccctctctg acaccccagg gtggccctga 1080
atccaggggc ctaggagcc cagctttaga atcaccgcgc tgggtactcg atggagcttg 1140
tctctgatgc agaacactcc tagcattctc tctcagggct cttttcattt gaatgacct 1200
gaggattgag ctcatgtagg cactgaaggc ttccacctct ccataaccg caaggccgat 1260
ctgccttcag ctcccagcaa gtgtggggca gcgcgggcca cagagtaggg tgcagggatg 1320
gggccccctgc agcaccagg gtctctggta tggagacagc agtgtggagt ctggaaactc 1380
agagtccttc tggctgccgc cgcggtttta ccatctggag agccaccacg ctgaagcctc 1440
ctccaccctg agcgcttggc tggcttcagg cctgtctcaa gatgcaagga gaggatacac 1500
caccatcctg ctggctgctc tgagtgtcac cccctgaaa gcagcacagg gtgccccctc 1560
catcctggca cccctactt ctccccagt ggatgcagaa tgtgctgggc caggtgctgg 1620
acgcgctgga atacctgcac catttggaca tcatccacag acccctttcg taagtcttg 1680
atggccccctg aagccctcaa cttctccttc agccagaaat cagacatctg gtccctgggc 1740
tgcatcattc tggacatgac cagctgctcc ttcattggatg gcacagaagc catgcatctg 1800
cggaagtccc tccgccagag cccaggcagc ctgaaggccg tcctgaagac aatggaggag 1860
aagcagatcc cggatgtgga aaccttcagg aatcttctgc ccttgatgct ccagatcgac 1920

ccctcggatc gaataacgat aaagtctgcc gtggcccccg gagctggtgg aggtggtggt 1980
cacgaccatg gagctacatg acagggtcct cgatgtccag ctgtgtgcct gctccctgct 2040
gtgacacctc ctgggccaag gtatcattgt gaacaaggcc cccttggaga aggtcccgga 2100
cctcatcagc caggtgttgg ccacctacc tgcggatggg gaaatggcag aagccagctg 2160
cggagtcttc tggctgctgt ccctgctggg ctgcatcaag gagcagcagt ttgaacaagt 2220
ggtggcgctg ctctgcaaa gcatccggct gtgccaggac agagccctgc tggatgaaca 2280
tgcctaccgg ggactggcca gcctggtgaa ggtgtcagag ctggcggcct tcaagggtgt 2340
ggtgcaggag gagggcggca gtggcctcag cctcatcaag gagacctacc agctccacag 2400
ggacgacccg gaggtggtgg agaacgtggg catgctgctg gtccacctgg ctctctatga 2460
ggagatcctg ccggagctgg tgtccagtag tatgaaggcc ctgctccagg agatcaagga 2520
gcgcttcacc tccagcctgg tgagtacag cagcgccttc agcaaaccag gcctccctcc 2580
aggtggaagc cccagctgg ggtgcaccac gtctggggga ctggaataga tgtttgtatg 2640
gaactgacct tgatctccac gtgtatagtt ttcaagactg ctctcctgcc tgcctattat 2700
cccatctcta tgactgggcc aaaatcaatc ttaaaccggga ggggtaatca gacctctcca 2760
aagagtttcc tgtccatgac tgctggattg agtcacatga gtaactgctc ctggacccgg 2820
ggactgtcca cgaaaactga cttgcctgct tcctccttcc aggaactggt ttcttgcgcg 2880
gaaaaagtgc tcttgaggct ggaggcagcc acctctccca gccactggg tggggaagca 2940
gctcagccct gatgcggggg agaagacaga taccacacag gccctccct ccacgtgtgc 3000
cctctccctg tccttccttt ccatgggcca ctgtttccct tggggtgggg ggaagggtca 3060
tccagcacca gaatgcgcat ctcacactcc tcttaggtga ctaataaaga ggcccaaggc 3120
c 3121

<210> 14

<211> 3257

<212> DNA

<213> Homo sapiens

<400> 14

tgaggaggcc caggaatctg tacttttaaat aaggccctgg agtctaagca gctgctgctg 60
tccctctaca ctaagccgct tgccagtgtg gagctcatga ttggcctgtg gtttcccagt 120
gattcacccg ggactgactc tgcctctgtg gatccggcta taacagcatc gtcattccaca 180
cctgtgatcc gctccagtgt cctccctccc accctatcct ctttgccttt tcatcctgtt 240
gtccgctgtt gtccttgggc tctggtctta ctgtcccatc cccagactgg tccttgtttg 300
tttcccttat ttaaggtgca ctcttcatct ttctcaggaa atatgggcca gctattggct 360
gagcatcctt cggggggctaa gcatgactat cctaattgca cagaggcttt ttctcataca 420
ttccttcctc tccccaattt cttcccatat ttgtttccaa agtggccata tcagtttctt 480
ctgctgtttg catatcttcc tgaaatgtta aaacaaccga acgacagcac agccctcatc 540
tctgtctagt attgcacaag aaagtgggct gttctctgtg ttgtagcatc tgttttcctc 600
ttgccaaatg agggcctggg aagaagaccc aggatccagt cctgatgctg ccaccaactc 660
acagtgtgac cctgggcatg tcacagctgc cccttggact tccagtgtg ccctgagaag 720
ctgcattgag agcagttagg agcacagaat ctgaaaccag cccaccagg caaaattcta 780
gctagcaact tggcaagtta cttaaactct gtctgcctta gtagtctcat ctatgaaatg 840
gagctaagaa taataattcc cctcttgaaa ttgagtgagt taagacatat aacattttta 900
gaagagtgcc tggcacagaa taatcagtgt catttccctc ccagtattaa aaaagagggt 960
tgctctaaat tagcacttcc tagaggatag tctatgagat tctttccac agggcactca 1020
gcaagagaaa aaggggagga gagagtcttg ggtacaggct tctggcagtg cagaaagcta 1080
ttaaacctta ttttgtctaa ttccagtctc tctcaaactc ccttgaccat agaatccatt 1140
tttcccagta cctcctagca tcctgtactt ctcccatgga acacttctga gaaacacccc 1200
tccagctttg acatgtatgg gaaaaaaggc actgggggta ttgagccctt actgtgtgtc 1260
actttgtgtg gatggctttg cagagttttc cccagctcc cctcagagga aactgaggcc 1320
agagagaggc cctggttaac aaatgtgttt tgttccaaat cctggactca gactactgtt 1380
atttactatc tcctttggtc ctaactgttg agagggggag gattggagac gactagaaca 1440
tgcaaggcct gtggtttagg gctgtcatct gcaagcagaa gtttgtggct ctggttttcca 1500
ggaggctaca ggacaccac atggcacagc ttctgacaa aatggcagct aatcaaactc 1560
gcctacctt ccacccccac actgacgttt gcctccaccg gaactccttt tctatgcatt 1620
tcatcttctc agatacagcg gtgcttctgt ttgatttctg gagtgtccac agtcctgctg 1680
gcatggccct ttcgggtgtg gtgctcctgc ttctggctgt actgtatgaa ggcattcaagg 1740

ttggcaaagc caagctgctc aaccaggtac tgggtgaacct gccaacctcc atcagccagc 1800
 agaccatcgc agagacagac ggggactctg caggctcaga ttcattccct gttggcagaa 1860
 cccaccacag gtggtatattg tgtcactttg gccagtctct aatccatgtc atccaggtgg 1920
 tcatcggcta cttcatcatg ctggccgtaa tgtcctacaa cacctggatt ttccttgggtg 1980
 tggctcttggg ctctgctgtg ggctactacc tagcttacc acttctcagc acagcttagc 2040
 tgggtgaggaa cgtgcaggca ctgaggctgg agggacatgg agccccctct tccagacact 2100
 atacttccaa ctgccctttc ttctgatggc tattcctcca ccttattctc agccccctgga 2160
 aactttgagc tgaagccagc acttgctccc tggagttcgg aagccattgc agcaaccttc 2220
 cttctcagcc agcctacata gggcccaggc atgggtcttgt gtcttaagac agctgctgtg 2280
 accaaaggga gaatggagat aacaggggtg gcagggttac tgagcccatg ataatgcttc 2340
 tctgtgactc aaaccaggaa tttccaaaga tttcaagcca gggagaaggg ttcttgggtga 2400
 tgcaggggcat ggaacctgga caccctcagc tctcctgctt tgtgccttat ctacaggagc 2460
 atcgcccat tggacttcctg acctcttctg tctttgaggg acagagacca agctagatcc 2520
 tttttctcac ctttctgcct ttggaacaca tgaagatcat ctctgtctatg gatcatgttg 2580
 acaaactaag ttttttttat ttttccatt gaactcctag ttggcaattt tgcacattca 2640
 tacaaaaaaa tttttaatga aatgatttca ttgattcatg atggatggca gaaactgctg 2700
 agacctat tccctttcttg gggagagaat aagtgcagc tgattaaagg cagagacaca 2760
 ggactgcttt caggtcctg gtttattctc tgattgatag actgagctcc ttccaccaga 2820
 aggcactgcc tgcaggaaga agatgatctg atggccgtgg gtgtctggga agctcttcgt 2880
 ggccctcaatg ccttccttta tctcatctt tcttctatgc agaacaaaaa gctgcatcta 2940
 ataatgttca atacttaata ttctctat ttttctatgc gcttactcgt aatgatctag 3000
 tggggaaaca tgattcattc acttaaaata ctgattaagc catggcaggt actgactgaa 3060
 gatgcaatcc aaccaaagcc attacat ttttgagttaga tgggactctc tggatagttg 3120
 aacctcttca ctttataaaa aaggaaagag agaaaatcac tgctgtatac taaatacctc 3180
 acagattaga tgaaaagatg gttgtaagct ttgggaatta aaaacaaaca aatacat ttt 3240
 agtaaatata aaaaaag 3257

<211> 3658

<212> DNA

<213> Homo sapiens

<400> 15

```
ggaacagtgt ggggcctgga ccgctgggta ggcgcgtcca gcggcctgag caggggaggg 60
taatgaggct gttacgcgcc ttctccgcat cttggcggga gcctgacgcc ccgcttcttc 120
cctaacgggg tgttccaccg gcgcctgccg aggcctaggc ctccgcagcc gccctccgtc 180
tcctcagccc cgacgctgcg cccgctttgt gcgcattttt ctctggggaa actgaggctc 240
cgagtgcgaa agtcagccga ggtcgccccg cccaggacag agaagggtg ggggtcggct 300
gagccgcggc attcccgggc cccgctaggg ctgcagggtc tcaggatggc agcctcggcg 360
caggtgtctg tgaccttga ggatgtggct gtgacattca cccaggagga gtggggacag 420
ttggatgcag cccagagaac cttgtatcag gaggtgatgc tggagacctg cggacttctc 480
atgtctctgg gctgtccttt gttcaaacca gagctgatct accagttgga tcacagacag 540
gagctatgga tggctacaaa agacctctcc caaagctcct atccaggatga caacacaaaa 600
cccaagacca cagagcctac cttttctcac ctggccttgc ctgaggaagt cttactccag 660
gaacaactga cacaaggagc ctcaaagaac tccaattag ggcaatcaa ggatcaggat 720
gggccatctg aatgcaaga agtccacttg aaaaatagga taggccccca gcgggggaag 780
ctgctggaga aatgagttc tgaacgtgat ggtttgggtt cagatgatgg tgtatgtaca 840
aagattacac agaaacaagt ttcaacagaa ggtgatctct atgaatgtga ttcacatgga 900
ccagttacag atgccttgat tcgcgaagag aaaaattcct ataatgtga ggaatgcggg 960
aaagtgttta aaaagaatgc cctccttggt cagcatgaac ggattcacac tcaagtgaag 1020
ccctatgaat gcacagagtg tgggaaaacc tttagcaaga gcactcatct tcttcagcac 1080
ctcatcatcc aactgggga gaagccctat aagtgcattg agtgtgggaa ggcttttaac 1140
cgcaggctgc acctcacacg gcaccagcgg attcacagtg gagagaagcc ttataagtgc 1200
agtgaatgtg gaaaggcctt caccaccgc tccacttttg tcttgcacac caggagccac 1260
actggagaaa aaccctttgt gtgcaaagag tgtggcaaag cctttcgaga taggccaggt 1320
ttcattcgac actacatcat ccacacggga gagaagccct atgagtgcac tgagtgtggg 1380
aaggccttca accgccggtc atacctcacg tggcaccaac agattcacac tggagtgaaa 1440
```

ccctttgaat gcaacgagtg tggaaaagct ttttgcgaga gtgcagacct cattcaacac 1500
tacattatcc aactgggga gaagccctat aagtgcattg agtgtgggaa ggcgttcaac 1560
cgtaggtcac acctcaagca gcatcaacgg attcacactg gggagaagcc ttatgaatgc 1620
agtgaatgtg gaaaggcctt caccactgc tccacttttg tcttgcataa aaggaccac 1680
acaggagaaa aaccctatga atgcaaagaa tgtggaaaag cctttagtga tagggcagac 1740
ctcattcgcc acctcagcat ccacactgga gagaaacct atgagtgcgt ggagtgtgga 1800
aaggccttca accgcagctc acacctcacg aggcaccaac agattcacac tggagagaaa 1860
ccctatgaat gcatccagtg tgggaaagcc ttttgcgga gcgcaaacct tattcgacac 1920
tccatcattc aactggaga gaagccgtat gaatgcagtg agtgtggaaa ggcttttaat 1980
cgcggtcat ccctcacaca tcatcaaagg attcatactg ggagaaacct taccattgta 2040
acagatgtgg gaagaccttt tatgactgca cagacttcag tcaacatcca ggaactttta 2100
ttagggaaaag agtttttgaa taccactt gaagaaaatc tgtggtgaaa gggaacatct 2160
taccatctgg ccattcacac tgaagagaaa cttcataagc atcctctctt tgagaaaacg 2220
tgtaatagat gatcatttgt cgtctaaaca atcaagttag aaattgaacc agcctggagt 2280
cttattctcc acctgagaat tcacccatga gagagacca gtggttgcta tgcacttagg 2340
aaaactttca gccacatctt tcttcttagt ttacagtga atattatctc agggatatct 2400
tttttttttt ttgagacgga gtctcgctct gtcgccagg ctggagtgca gtggcgggat 2460
ctcggtcac tgcaagctcc gcctcccggg ttcacgccat tctctgcct cagcctcccg 2520
gggtggctggg actacaggca cccgccacta cgcctggcta gttttttgta ttttagtag 2580
agacggggtt tcaccgtttt ggccgggatg gtctcgatct cctgacctcg tgatccgcc 2640
gcctcagcct ccaaagtgc tgggattaca ggctgagcc accgcgccg gccatctcag 2700
ggatatttaa acaaaggaag aggaaatata gggaagagaa agaaaccaca gcttctttta 2760
gacttgtctg acaagcctat ggcaattttg tcgtcccttt gttattttat gtgggagagg 2820
gaaatgtttc agaaacgaaa gatctcatcc cctttatttt tctgtgtatg cattcactgc 2880
tgcccagtgt cagaacagtt atacaaacga gggcttgaaa acattttgtg aaaacttgct 2940
tagtcctgca acgttctcca ctctcgaggt gcaggagcat acatctttat aaaaaatatg 3000
gaggcttgag ttatgaaaca cttttataag gagataagga tgtacatatg tttggacaca 3060
ttgtccactt cagatgcttg tttaaaaaaa tgtttcacca tgtgtcttta cccaatacac 3120
atactttttc ttgatgtgga tttacttttt tttttttttt ttaagttttt ctgggatagg 3180

gatgatagtg taattgttca tgcccacagt aatcctggga atcttagttt tttcctatga 3240
 ttgtggtttc tcagcttctt tacctctcct ttgagttagg tccacctcag ttttctaaag 3300
 gtcaaattaa ggcccccca aaaatacagt atgtagctga attcctttga aagatcatga 3360
 taaactcaga agaacaaca tagaaggaat gtcattttta atctttttaa agagtgtgtg 3420
 aaaattcact ttaactggaa ttttaacagg caagccggca tgcatatatg tagggatgtg 3480
 tgactttcac agttgtcacc atggacttag gcttgattaa aatgctacac atttagagcg 3540
 ttttgaatt ctttttttat atttttaaga ctagattttg agccatgtga tagttttacc 3600
 tcatattaaa tatattgtac agaagtaaaa ataatcatgt attcactttg aactcctc 3658

<210> 16

<211> 2800

<212> DNA

<213> Homo sapiens

<400> 16

gatcatgctc aggtgggtcca tggagttatt tgcagcatTT cttgctcctg gaagtggggT 60
 tgaggggcccg aggagaagca gaggtgctc ggcagagatc ccggaacact ttatccctct 120
 ctcttgccca ggttttatgc cgctgagata atgtgtggac tgcagtttct acacagcaag 180
 ggcatcattt acaggtgcgg ggggtgagggc agcgggggct cttgggaggg gaggtccag 240
 ccccatcata tcttctgaaa tgctccaagc aggatcccc caactccagt tccttctctg 300
 ctggaaatca atctttagct gggatattgc ctatgcctcc caccctgggtg ttagcattcc 360
 tgtggccacc atttgaggct ggctcagctt caggaaactag gaactcactg ctcatatttg 420
 acaccatcgc ccacccatc cttggagccg tcagtctcgc tctgaagccc catcgagtca 480
 tctaagcttc tgccagaagg gactccttcc ataaagtctc atctaataag ttatgattcc 540
 atcctatggT gggaactgta caccaccaa aaaggctgag gtcaggctct gtatactgat 600
 gtgggataat gccaagatat gttgttaagt ggaaaaacaa cctgtgggaa agtgtatatc 660
 gaatattagc attcagcctt taaaacgtgg gattataggc gtgagccatc gtgcctggcc 720
 cgcacttata aattagatat aactataatg tataataggg tcattgtgaa gatcaaatgg 780

aataagatat gtaaagctgt attttaagtg ccttccactg gtagccgttc tgtgtaagaa 840
gtctgttatac attattacac atgctttgtg tgtatgtaca ggtgtctgtt ttgaggcgtg 900
tccatgtgca tgtgtaagag agggcattta ggagccgttt gctatcccag ggcttggaaag 960
aactgcccc gcttctcata ttgcgatgcc ctgtgcccgc taacttctgc ccacctccct 1020
gcccactagg actccttctt tctaagcccc tgcagattgg cctgtcctct ccgccccgtc 1080
ctctccaggc tcctccttcg agggctggca ggaggaagac tcaagcgctg ggcctctgcg 1140
gggtgagggt gtgggcggtc aggagaacgg tggctttgtg tagagggcta gactggctcg 1200
caggcaccag ctcatagact ctgccccctc cgggtgtttc ccttccccct cctgggcctg 1260
tgctcttca agaatgtaac gctgccatcc acaagaaatg catcgacaag atcatcgga 1320
gatgcactgg caccgcggcc aacagccggg acgctatatt ccagaaagaa cgcttcaaca 1380
tcgacatgcc gcaccgcttc aaggttcaca actacatgag cccaccttc cgtgaagatg 1440
aaggaggcgc tcagcacaga gcgtgggaaa aacttggtgc agaagaagcc gaccatgtat 1500
cctgagtgga agtcgacgtt cgatgcccac atctatgagg ggcgcgtcat ccagattgtg 1560
ctaatacggg cagcagagga gccagtgtct gaggtgaccg tgggtgtgtc ggtgctggcc 1620
gagcgtgca agaagaacaa tggcaaggct gagttctggc tggacctgca gcctcaggcc 1680
aaggtgttga tgtctgttca gtatttcctg gaggacgtgg attgcaaaca gtctatgcgc 1740
agtgaggacg aggccaagtt cccaacgatg aaccgccgcg gagccatcaa acaggccaaa 1800
atccactaca tcaagaacca tgagtttata gccaccttct ttgggcaacc caccttctgt 1860
tctgtgtgca aagactttgt ctggggcctc aacaagcaag gctacaaatg caggcgtaag 1920
tgtctccaca ggccagtttg tacgtatgta cccatgtgtg ctcacgtgtg ccagtgcctg 1980
tgtgtatgcc agtgcctgtg tgcgctcaga gagtgtatgc acgtgagttt tcccagtgt 2040
gctacagcag ctgagttcag tgagtgtgg ggctggcttg ctgctgctct ggaaaaccag 2100
ttctgcgcat ctcttcccag ctccgcattc agtgacgggg agttggtgga tggcaattgg 2160
gcacagtaga gtattcacgc agaaaccagc ccaggccatg caccaggcct cttttccccg 2220
gagaggcaat gataggtcct ttactggcac atagtggctg catgtgcttc tgtgtagatg 2280
tgggtgtctc tgagtacaca ggtgtgagca cactctgtcc ctgagggtgt atacacgagg 2340
cgccgcaggc tggagccaga ctgcctgggt ttgaatccca gctctaccac ttctagcta 2400
tagggcctta ggcaaagtta cttaacctct ctgtgcctca gtttctcat ctataatttt 2460
tttttttttg agactgagtc tcattctgtt gcccaagcta gactgcagtt gtgcgatctc 2520

agctcactgt aacctccctc tccaaggttc aagcgattct cttgccttag ctccttgagt 2580
agctgggatt acaggcgtgc accagcctga ccaacatggt gaaaccctgt ctctactaaa 2640
aaatacaaaa gttagctggg catggtggtg gatgcctgta atctcagcta ctcaggaggc 2700
tgaggcagga gaattgcttg aaccgggag gcgagggttg cagtgagctg agatcatgcc 2760
accgcactcc agcctgtgtg acagaatgag actgtgtctc 2800

<210> 17

<211> 5202

<212> DNA

<213> Homo sapiens

<400> 17

gtctccagga ctgacgtca ggctcctctc tcgccttagc ccaacttgct ttcccgcctc 60
gcaaaactccg gtttccctcc actcccaact cttttcacta cacgtttccc ctctctatc 120
tcccacgcca cgaaccccgga tcccagact cctctctccc gccctcctcc ttctctctc 180
ctcccttcaa ctcttcatcc gcttccacct cagactctgc gcgcacccaa ttcagtcgcc 240
cgctcccgtt cggctcctcg aagccatggc gggacctggg ggctggaggg acagggaggt 300
cacgatctg ggccacctgc cggatccaac tggaatatc tcactagata aaaccattgg 360
ccttggtact tatggcagaa tctatgtggg acttcatgag aagactggtg catttacagc 420
tgttaaagt atgaacgtc gtaagacccc ttacctgaa ataggaaggc gagtgagagt 480
gaataaatat caaaaatctg ttgggtggag atacagtgat gaggaagagg atctcaggac 540
tgaactcaac cttctgagga agtactcttt ccacaaaaac attgtgtcct tctatggagc 600
atttttcaag ctgagtcctc ctggtcagcg gcaccaactt tggatgggtga tggagtatat 660
tgcggcaggt tcggtcactg atgtagttag aatgaccagt aatcagagtt taaaagaaga 720
ttggattgct tatacttgcc gagaaatcct tcagggttta gctcaccttc acgcacaccg 780
agtaattcac cgggacatca aaggtcagaa tgtgctgctg actcataatg ctgaagtaaa 840
actggttgat tttggagtga gtgcccaggt gagcagaact aatggaagaa ggaatagttt 900
cattgggaca ccatactgga tggcacctga ggtgattgac tgtgatgagg acccaagacg 960

ctcctatgat tacagaagtg atgtgtggtc tgtgggaatt actgccattg aaatggctga 1020
aggagcccct cctctgtgta accttcaacc cttggaagct ctcttcgtta ttttgcggga 1080
atctgctccc acagtcaaatt ccagcggatg gtcccgttaag ttccacaatt tcatggaaaa 1140
gtgtacgata aaaaatttcc tgtttcgtcc tacttctgca aacatgcttc aacacccatt 1200
tgttcgggat ataaaaaatg aacgacatgt tgttgagtca ttaacaaggc atcttactgg 1260
aatcattaaa aaaagacaga aaaaaggatt tgcagtgagc caagatcatg cactgcact 1320
ccagcccggg tgacagaaca agacgtgcc tcaaaaaaaa taaaaataaa agtaccagga 1380
tggattttca ccaagtccag aagaagcccc tgggaagtcc taactttata tattatatta 1440
catattccac atttgatttt gcggaccata ggcagagttg agagtaacaa tggaatgtaa 1500
aagggaanaac ttgaaaaagc aggcacatcctc agatcagctg tgggattgat tgtagtgggg 1560
cttggtgtat ctatacatgt atcatgtttt gaggcaagga gcagcaagaa ttaatttatt 1620
tacctatggt aatgaagtta ttcagccaac ttacagagg ccagccagtt tcaaatgggt 1680
tcaggaaaac cacacatcta tatggtaatt tgctaatttc cagaattgca aaatttccaa 1740
aatgttacct gtaggaaaag ttccatttat atggaagtta cactaaagat ataaatatag 1800
ttgccctgaa cattttaata taggaaattc aaaataaaac tcttgtttgt aatcttctat 1860
atttcagtag tcacagtaca gctgatgata ggaaagaagc tgaatgtttt caatgtgttt 1920
agtgtaaaac acagaattaa gaatttgcct tctcattgaa aagaccaaatt atcaacctca 1980
agccataaag tccttacatc tagagccac agctgtaagc catgctcacc actatttaga 2040
gtttctcttg tctcattctc cttttgtccc attctctctt cttcctcttt agaattgcag 2100
aaatcaagtt gctaaatttt ctctgttcag agaaattaat ctgcatgcaa tagctatatt 2160
gtgctgcaa agtgggccta gaaactagaa tatggtctac atcacaacc aagtactttt 2220
gaagggtac cttcttgccc ttgtaattca ctccacatcc taagaagcag agggaacaag 2280
gtgattaagt ggataacatg cttatctgct gactgtttat gcattatgtt tttctctttt 2340
gtaaaaaaag gaataccttt gatctttgaa agagaagaag ctattaagga acagtacacc 2400
gtgagaagat tcagaggacc ctcttgact cacgagcttc tgagattgcc aaccagcagc 2460
agatgcagac cacttagagt cctgcatggg gaacctctc agccaagggtg gctacctgat 2520
cgagaagagc cacaggtcca ggcacttcag cagctacagg gagcagccag ggtattcatg 2580
ccactgcagg ctctggacag tgcacctaag cctctaaagg ggcaggctca ggcacctcaa 2640
cgactacaag gggcagctcg ggtgttcatg ccactacagg ctcagggtgaa ggctaaggcc 2700

tctaaacctc tacaaatgca gattaaggca cctccacgac tacggagggc agccagggtg 2760
ctcatgccac tacaggcaca ggtagggca cccaggcttc tgcaggtaca gtcccaggta 2820
tccaaaaagc agcaggccca gaccagaca tcagaaccac aagatttgga ccaggtagca 2880
gaggaatttc aggggtcaaga tcaggtagcc gaacaacaaa ggcagggccca ggcccctgaa 2940
caacagcaga ggcacaacca ggtgcctgaa caagagctgg agcagaacca ggcacctgaa 3000
cagccagagg tacaggaaca ggctgccgag cctgcacagg cagagactga ggcagaggaa 3060
cctgagtcac tacgagtaaa tgcccaggta tttctgcccc tgctatcaca agatcacat 3120
gtgctgttgc cactacattt ggatactcag gtgctcattc cagtagaggg gcaaactgaa 3180
ggatcacctc aggcacaggc ttggacacta gaacccccac aggcaattgg ctgagttcaa 3240
gcactgatag agggactatc aagagacttg cttcgggcac caaactcaaa taactcaaag 3300
ccacttggtc cgttgcaaac cctgatggaa aatctgtcat caaatagggt ttactcaca 3360
ccagaacagg cacgggagaa aaaatcaaaa gtttctactc tgaggcaagc actggcaaaa 3420
agactatcac caaagagggt cagggcaaag tcacatgga gacctgaaa gcttgaactc 3480
tcggatttag aagcccgag gcaaaggcgc caacgcagat gggaagatat ctttaatcag 3540
catgaggaag aattgagaca agttgataaa gacaaagaag atgaatcatc agacaatgat 3600
gaagtatttc attcgattca ggctgaagtc cagatagagc cattgaagcc atacatttca 3660
aatcctaaaa aaattgaggt tcaagagaga tctccttctg tgcctaaca ccaggatcat 3720
gcacatcatg tcaagttctc ttcaagcgtt cctcagcgtt ctcttttgga acaagctcag 3780
aagcccattg acatcagaca aaggagttcg caaaatcgtc aaaattggct ggcagcatca 3840
gaatcttctt ctgaggaaga aagtcctgtg actggaagga ggtctcagtc atcaccacct 3900
tattctacta ttgatcagaa gttgctgggt gacatccatg ttccagatgg atttaaagta 3960
ggaaaaatat caccctgt atacttgaca aacgaatggg taggctataa tgcactctct 4020
gaaatcttcc ggaatgattg gttaactccg gcacctgtca ttcagccacc tgaagaggat 4080
ggtgattatg ttgaactcta tgatgccagt gctgatactg atggtgatga tgatgatgag 4140
tctaatagata cttttgaaga tacctatgat catgccaatg gcaatgatga cttggataac 4200
caggttgatc aggctaata tgtttgtaaa gaccatgatg atgacaacaa taagtttggt 4260
gatgatgtaa ataataatta ttatgaggcg cctagttgtc caagggaag ctatggcaga 4320
gatggaagct gcaagcaaga tggttatgat ggaagtcgtg gaaaagagga agcctacaga 4380
ggctatggaa gccatacagc caatagaagc catggaggaa gtgcagccag tgaggacaat 4440

gcagccattg gagatcagga agaacatgca gccaatatag gcagtgaaag aagaggcagt 4500
 gaggggtgatg gaggtaaggg agtcgttcga accagtgaag agagtggagc ccttggactc 4560
 aatggagaag aaaattgctc agagacagat ggtccaggat tgaagagacc tgcgtctcag 4620
 gactttgaat atctacagga ggagccagggt ggtggaaatg aggcctcaaa tgccattgac 4680
 tcagggtgctg caccgtcagc acctgatcat gagagtgaca ataaggacat atcagaatca 4740
 tcaacacaat cagatttttc tgccaatcac tcatctcctt ccaaagggtc tgggatgtct 4800
 gctgatgcta actttgccag tgccatctta tacgctggat tcgtagaagt acctgaggaa 4860
 tcacctaagc aaccctctga agtcaatgtt aaccactct atgtctctcc tgcattgtaa 4920
 aaaccactaa tccacatgta tgaaaaggag ttacttctg agatctgctg tggttctttg 4980
 tggggagtca atttgctgtt gggaacccga tctaattctat atctgatgga cagaagtgga 5040
 aaggctgaca ttactaaact tataaggcga agaccattcc gccagattca agtccttagag 5100
 ccaactcaatt tgctgattac catctcagggt cataagaaca gacttcgggt gtatcatctg 5160
 acctggctga ggaacaagat tttgaataat gatccagaaa gt 5202

<210> 18

<211> 3312

<212> DNA

<213> Homo sapiens

<400> 18

acaaagagtt ggcagatcac ggatggaggg cagcatctcc caacagcctg ggcggccgct 60
 gagaccaga gaaccaagg actcccctgg gctcatccag cagcctctgc tcccaggag 120
 agagggtgctg aagtccacga agagggtctc gctctgtcac acaggctgga gtgcagtgg 180
 gtgatcttgg ctcatcgtaa cctccacctc ccgggttcaa gtgattctca tgcctcagcc 240
 tcccagtag ctgggattac aggtggtgac ttccaagagt gactccgtcg gaggaaaatg 300
 actccccagt cgctgctgca gacgacactg ttcctgctga gtctgctctt cctgggtccaa 360
 ggtgcccacg gcaggggcca cagggaagac tttcgcttct gcagccagca gttgcagagc 420
 ctggagtcga aactgacctc tgtgagattc atggggggaca tggtgtcctt cgaggaggac 480

cggatcaacg ccacggtgtg gaagctccag cccacagccg gcctccagga cctgcacatc 540
cactcccggc aggaggagga gcagagcgag atcatggagt actcgggtgct gctgcctcga 600
acactcttcc agaggacgaa aggccggagg ggggaggctg agaagagact cctcctggtg 660
gacttcagca gccaagccct gttccaggac aagaattcca gccacgtcct gggtgagaag 720
gtcttgggga ttgtgttaca gaacaccaaa gtagccaacc tcacggagcc cgtgggtgctc 780
accttccagc accagctaca gccgaagaat gtgactctgc aatgtgtgtt ctgggttgaa 840
gaccccatat tgagcagccc ggggcatttg agcagtgtg ggtgtgagac cgtcaggaga 900
gaaacccaaa catcctgctt ctgcaaccac ttgacctact ttgcagtgtc gatgggtctc 960
tcggtggagg tggacgccgt gcacaagcac tacctgagcc tcctctccta cgtgggtgtg 1020
gtcgtctctg ccctggcctg ccttgtcacc attgccgcct acctctgtc cagggtgccc 1080
ctgccgtgca ggaggaaacc tcgggactac accatcaagg tgcacatgaa cctgctgctg 1140
gccgtcttcc tgctggacac gagcttctctg ctacgcgagc cgggtggcct gacaggctct 1200
gaggctggct gccgagccag tgccatcttc ctgcacttct ccctgtcac ctgcctttcc 1260
tggatgggccc tcgaggggta caacctctac cgactcgtgg tggaggtctt tggcacctat 1320
gtccctggct acctactcaa gctgagcgcc atgggctggg gcttcccat ctttctggtg 1380
acgctggtgg ccctggtgga tgtggacaac tatggcccca tcactttggc tgtgcatagg 1440
actccagagg gcgtcatcta cccttccagg tgctggatcc gggactccct ggtcagctac 1500
atcaccaacc tgggcctctt cagcctggtg tttctgttca acatggccat gctagccacc 1560
atggtggtgc agatcctgcg gctgcgcccc cacacccaaa agtggtcaca tgtgtgaca 1620
ctgctgggccc tcagcctggt ccttggcctg ccctgggcct tgatcttctt ctcctttgct 1680
tctggcacct tccagcttgt cgtcctctac cttttcagca tcatcacctc cttccaagge 1740
ttcctcatct tcacttggtg ctggtccatg cggctgcagg cccgggggtg cccctccct 1800
ctgaagagca actcagacag cgccaggctc cccatcagct cgggcagcac ctcgtccagc 1860
cgcatctagg cctccagccc acctgcccct gtgatgaagc agagattcgg cctcgtcgca 1920
cactgcctgt ggcccccgag cccggcccag ccccaggcca gtcagccgca gactttggaa 1980
agcccaacga ccatggagag atgggcccgt gccatggtgg acggactccc gggctgggct 2040
tttgaattgg ccttggggac tactcggtc tcactcagct cccacgggac tcagaagtgc 2100
gccgccatgc tgcctagggt actgtcccca catctgtccc aaccagctg gaggcctggt 2160
ctctccttac aaccctggg cccagccctc attgctgggg gccaggcctt ggatcttgag 2220

ggtctggcac atccttaate ctgtgccctt gcctgggaca gaaatgaggc tccagttgct 2280
 ctgtctctcg tggtcaccct gagggcactc tgcatcctct gtcattttta cctcaggtgg 2340
 caccaggggc gaatggggcc cagggcagac cttcagggcc agagccctgg cggaggagag 2400
 gccctttgcc aggagcacag cagcagctcg cctacctctg agcccaggcc ccctccctcc 2460
 ctcagcccc cagtcctccc tccatcttcc ctggggttct cctcctctcc cagggcctcc 2520
 ttgctccttc gttcacagct ggggggtccc gattccaatg ctgttttttg gggagtgggt 2580
 tccaggagct gcctgggtgc tgctgtaaat gtttgtctac tgcacaagcc tcggcctgcc 2640
 cctgagccag gctcggtacc gatgcgtggg ctgggctagg tccctctgtc catctgggcc 2700
 tttgtatgag ctgcattgcc cttgctcacc ctgaccaagc acacgcctca gaggggccct 2760
 cagcctctcc tgaagccctc ttgtggcaag aactgtggac catgccagtc ccgtctgggt 2820
 tccatccac cactccaagg actgagactg acctcctctg gtgacactgg cctagggcct 2880
 gacactctcc taagaggttc tctccaagcc cccaaatagc tccaggcgcc ctcgccgcc 2940
 catcatgggt aattctgtcc aacaaacaca cacgggtaga ttgctggcct gttgtaggtg 3000
 gtagggacac agatgaccga cctgggtcact cctcctgcca acattcagtc tggtagtgga 3060
 ggcggtgcgtg aagcaagaac tcctggagct acagggacag ggagccatca ttcctgcctg 3120
 ggaatcctgg aagacttctc gcaggagtca gcgttcaatc ttgacctga agatgggaag 3180
 gatgttcttt ttacgtacca attcttttgt cttttgatat taaaagaag tacatgttca 3240
 ttgtagagaa tttggaaact gtagaagaga atcaagaaga aaaataaaaa tcggctgttg 3300
 taatccccta gc 3312

<210> 19

<211> 1864

<212> DNA

<213> Homo sapiens

<400> 19

atcagaccaa cgtgaattga atcacagcag ttcagaaaga ggaaactgga ttcaggactt 60
 gaacgcaact cccaattgca gaaaattggc aacgtctctg aagagccctt gcttttgcct 120

ggacccccag catcatggtt tccatttca tggggtctct cagtgtcctg tgtttccttc 180
tgctgcttgg attccagttc gtctgcccac agccctccac tcaacacagg aaggtcccgc 240
agcggatggc ggcggagggc gccccgagg acgacggcgg cggcggcgcc ccgggagtgt 300
ggggcgccctg gggcccctgg tcggcctgct cgcgtagctg cagcggcggc gtgatggagc 360
agacgcggcc ctgcctgccc cgctcctacc gcctgcgcgg cggccagcgg cctggcgccc 420
ctgcgcgcgc cttcgcggac cacgcggtgt cggcgggtgcg cacgtcgggtg cactgcacc 480
ggagccgcga cgagacgcca gcgctggccg gtacggacgc cagccgccag ggccccacgg 540
tgctgcgagg cagccggcac ccacagcccc agggcctcga agtcactggg gacagaagga 600
gcaggaccgg tgggtaccatt ggccctggca agtatggcta tggtaaggcc ccatatatct 660
taccactgca gacagacact gcacacacgc cacagaggct ccggagacag aagctctcat 720
cccgccattc caggtcccag ggagcatctt ctgctaggca tggctacagt tcaccagccc 780
accagggtccc ccaacatggg cttttgtacc aaagtgcagc tggccctcgc tctggactgc 840
aggctgcgga ggccccatc taccagctac ctttgacca tgatcaaggc taccctgcag 900
cttcaagtct ctttcacagc ccagaaacaa gcaacaacca cgggtgtgggg acccatgggg 960
caactcagag cttctctcag cctgcccgat ctacagcaat ctcatgcac ggggcctatc 1020
ggcagtacaa gctgtgcaac accaacgtat gtccagaaag cagtagaagt atccgggagg 1080
tacagtgtgc atcctacaac aacaagccat tcatgggccg gttttatgag tgggaaccat 1140
ttgcagaagt aaaaggcaat cgcaaatgtg agttgaactg ccaggcaatg ggctaccgct 1200
tctatgtacg gcaagctgag aaagtcacgc atggcaccac ctgtgaccag aacggcacgg 1260
ccatctgtgt gtctgggcag tgcaagagca ttggctgtga tgactactta ggctccgaca 1320
aagtcgtgga caaatgtggg gtgtgtggag gagacaacac gggctgtcag gttgtgtcgg 1380
gcgtgtttaa gcatgccctc accagcctgg gctaccaccg cgtcgtggag atttccgagg 1440
gagccacgaa aatcaacatc acggagatgt acaagagcaa caactatttg gccctgagaa 1500
gtcgttcttg acgtccatc atcaatggga actgggcaat tgatcgacca ggaaaatacg 1560
agggcggagg gaccatgttc acctacaagc gtccaaatga gatttcgagc actgccggag 1620
agtccttttt ggcggaaggc cccaccaacg agatcttggg tgtctacatg atacaccagc 1680
agccaaaccc aggcgtgcac tacgagtacg tgatcatggg gaccaacgcc atcagcccc 1740
aggtgccacc ccacaggaga ccagggaac cttcaatgg ccagatgggtg acagaaggca 1800
ggagccagga ggaggagaa cagaaaggga ggaacgagga gaaggaagac ttgcgtgggg 1860

aggc

1864

<210> 20

<211> 3119

<212> DNA

<213> Homo sapiens

<400> 20

acacctcccc	gcgccgccgc	cgccaccgcc	cgcactccgc	cgcctctgcc	cgcaaccgct	60
gagccatcca	tgggggtcgc	gggccgcaac	cgtcccgggg	cggcctgggc	ggtgctgctg	120
ctgctgctgc	tgctgccgcc	actgctgctg	ctggcggggg	ccgtcccgcc	gggtcggggc	180
cgtgccgcgg	ggccgcagga	ggatgtagat	gagtgtgccc	aagggctaga	tgactgcat	240
gccgacgcc	tgtgtcagaa	cacaccacc	tcctacaagt	gctcctgcaa	gcctggctac	300
caaggggaag	gcaggcagt	tgaggacatc	gatgaatgtg	gaaatgagct	caatggaggc	360
tgtgtccatg	actgtttgaa	tattccaggc	aattatcggt	gcacttggtt	tgatggcttc	420
atgttggctc	atgacgggtc	taattgtctt	gatgtggacg	agtgcctgga	gaacaatggc	480
ggctgccagc	atacctgtgt	caacgtcatg	gggagctatg	agtgtctgct	caaggagggg	540
tttttcctga	gtgacaatca	gcacacctgc	attcaccgct	cgggaagagg	cctgagctgc	600
atgaataagg	atcacggctg	tagtcacatc	tgcaaggagg	ccccaagggg	cagcgtcgcc	660
tgtgagtgca	ggcctgggtt	tgagctggcc	aagaaccaga	gagactgcat	cttgacctgt	720
aaccatggga	acggtgggtg	ccagcactcc	tgtgacgata	cagccgatgg	cccagagtgc	780
agctgccatc	cacagtacaa	gatgcacaca	gatgggagga	gctgccttga	gcgagaggac	840
actgtcctgg	aggtgacaga	gagcaacacc	acatcagtgg	tggatgggga	taaacgggtg	900
aaacggcagc	tgctcatgga	aacgtgtgct	gtcaacaatg	gaggctgtga	ccgcacctgt	960
aaggatactt	cgacaggtgt	ccactgcagt	tgtcctgttg	gattcactct	ccagttggat	1020
gggaagacat	gtaaagatat	tgatgagtgc	cagaccgca	atggaggttg	tgatcatttc	1080
tgcaaaaaca	tcgtgggcag	ttttgactgc	ggctgcaaga	aaggatttaa	attattaaca	1140
gatgagaagt	cttgccaaga	tgtggatgag	tgctcttttg	ataggacctg	tgaccacagc	1200

tgcatacaacc accctggcac atttgcttgt gcttgcaacc gagggtagac cctgtatggc 1260
ttcaccact gtggagacac caatgagtgc agcatcaaca acggaggctg tcagcaggctc 1320
tgtgtgaaca cagtgggcag ctatgaatgc cagtgccacc ctgggtacaa gctccactgg 1380
aataaaaaag actgtgtggc ttcttgtgac ctgagctgca tcgtaaagcg aaccgagaag 1440
cggctccgta aagccatccg cacgctcaga aaggccgtcc acaggagca gtttcacctc 1500
cagctctcag gcatgaacct cgacgtggct aaaaagcctc ccagaacatc tgaacgccag 1560
gcagagtcct gtggagtggg ccagggtcat gcagaaaacc aatgtgtcag ttgcagggtc 1620
gggacctatt atgatggagc acgagaacgc tgcattttat gtccaaatgg aaccttccaa 1680
aatgaggaag gacaaatgac ttgtgaacca tgcccaagac caggaaattc tggggccctg 1740
aagaccccag aagcttgga tatgtctgaa tgtggaggctc tgtgtcaacc tgggtgaatat 1800
tctgcagatg gctttgcacc ttgccagctc tgtgccctgg gcacgttcca gcctgaagct 1860
ggtcgaactt cctgcttccc ctgtggagga ggccttgcca ccaaactca gggagctact 1920
tcctttcagg actgtgaaac cagagttcaa tgttcacctg gacatttcta caacaccacc 1980
actcaccgat gtattcgttg ccagtgaggc acataccagc ctgaatttgg aaaaaataat 2040
tgtgtttctt gccaggaaa tactacgact gactttgatg gctccacaaa cataacccaa 2100
ttctgtgaca acatatgaaa cctgccagac ctacgaacgc cccatcgcct tcacctccag 2160
gtcaaagaag ctgtggattc agttcaagtc caatgaaggg aacagcgcta gagggttcca 2220
ggccccatac gtgacatatg atgaggacta ccaggaaactc attgaagaca tagttcgaga 2280
tggcaggctc tatgcatctg agaaccatca ggaaatactt aaggataaga aacttatcaa 2340
ggctctgttt gatgtcctgg cccatcctca gaactatttc aagtacacag cccaggagtc 2400
ccgagagatg tttccaagat cgttcacccg attgctacgt tccaaagtgt ccaggttttt 2460
gagaccttac aatgactca gccacgtgc cactcaatac aatgttctg ctatagggtt 2520
gggtgggacag agctgtcttc cttctgcatg tcagcacagt cgggtattgc tgcctcccgt 2580
atcagtgact cattagagtt caatttttat agataatata gatattttgg taaattgaac 2640
ttggtttttc tttcccagca tcgtggatgt agactgagaa tggctttgag tggcatcagc 2700
ttctcactgc tgtgggcgga tgtcttggat agatcacggg ctggctgagc tggactttgg 2760
tcagcctagg tgagactcac ctgtccttct ggggtcctac tcctcctcaa ggagtctgta 2820
gtggaaagga ggccacagaa taagctgctt attctgaaac ttcagcttcc tctagcccgg 2880
ccctctctaa gggagccctc tgcactcgtg tgcaggctct gaccaggcag aacaggcaag 2940

aggggaggga aggagacccc tgcaggctcc ctccacccac cttgagacct gggaggactc 3000
agttttctcca cagccttctc cagcctgtgt gatacaagtt tgatcccagg aacttgagtt 3060
ctaagcagtg ctcgtgaaaa aaaaaagcag aaagaattag aaataaataa aaactaagc 3119

<210> 21

<211> 2356

<212> DNA

<213> Homo sapiens

<400> 21

atataaaatg ccgggttaag cggcaactca gactcaggat cccgctcacg acatggcctc 60
gggcgctcag ctcccgccgc agccgtcgag ctccagaggtc agcgccgtcc agagcccagg 120
cgggcgtccc ggccgcccgtc tggaggaaac agccctgggc gttcctctcc cgccgtctcc 180
ggggggaggcc cctctgcccc gaagcaaccg gagcagggtc cctgggacct gccagcccgg 240
agcggcctcc ctccacgcgg cgtccgcagc agtccccgtg cggccccggc gcggtacggc 300
gccagccggg aaaaccgcag acgcgggtccc cgccgcccgc ccagagcaag ctccgcggcc 360
ggctccacag tcccgcgaagc cgcgcaacct ggaaggcgac ctggacgagc gccggctgct 420
ctgccacttg cagctggccc aggaccgcga ggccgcgctg tggcggggcg gcaaacccca 480
ggatgaaatc tgcgacgcct tcgaggaagt cgtgctgtgg ctccctgcggc ttcagaacac 540
cttttacttc tcccagtcca cttttaacct ggccctcacc atctttggcc gcctcctgat 600
ttcagtgaag gtaaaagaga aatacctgca ttgcgccaca attacttcct tgaggctcgc 660
tgcaaaagtt aatgaagaag aggagtttat tccacaagta aaagacttca caaagcacta 720
tggctctgac tattccccga atgagctgct gaggatggag ctggctattc tggacagact 780
gcactgggac ctctatattg ggacgccgct ggacttcttg actatattcc atgccctggt 840
ggtcctgagc tggcccatg tgttggagct gctgcctcag aggaatcctt ccctccacgt 900
cgcatccctg accaggcagc tgcagcactg tatggcgggc caccagctgc tgcagttcaa 960
gggctccaca ctggccttgg tcatcatcac cttagagctg gagaggctca tgcccggctg 1020
gtgtgctcct atatctgac tgctaaagaa agcacagggt ggtgatatgc agtacagctg 1080

ctgcaaggaa cttgtaatgc agcaactgag aagtcttcag tcatcctcct gcacagacaa 1140
cttttgtgtca cctgccaact agcccctctg cctccacccc ggggctttca gagcatagtg 1200
tgaaacctcc ttgcttggaac taccatgagt tctttggctt gttatgaatc ctgtaaaaag 1260
ggaaggtggc tctggaagag caactgagaa aaagtccca actgagccct tggaaaaaaa 1320
ataaagggga gaggggaaag gcaggctgag gtcagtagag gtcagggtgg tctgatagac 1380
tatgaccctg tcttcccttt ttctcctctg ggccctgaagc caggagatgaat gaatgaatgt 1440
tcaaattggca gcttgtttta ccttccttct ccagcaaggg ttggcattgg cccctgggag 1500
cgctggaata tgaaaccaag aggcaggcct ggctcagggc tgaagggtg gggctagctc 1560
tgaccagtgc ttgctgctct acctgccaga gctggccagg tccttccaac ttcccaggct 1620
tatcccaaac acatatgtgg tgcttccagt cccaactacc ctgcctgcca ctcctcatat 1680
gggacagtct ctgcagatgc atttacccea ccatggctat gacttatgtc ttccctttac 1740
aaggactgga ttataaagta tgctttgttt aatgtaactt gagcaagctt tagacatata 1800
tttcctttgt tactcttaag caaagtgggt tcagttatta gtttaactag agggattttg 1860
tgctaataaa tgggtattat aggctaacac ctgtgatggg aggcacagct ctgaatgctt 1920
tctgttggga tgagcagggt ctgaggccaa gcctgctcca accacagtgc tgccaggagg 1980
aggagccggc tgcctctctc atgctgtgct accacagagc cattctctaa gcaggggagt 2040
gcagctgaca aagatgccag ctggggctgt gtaagtgtg ttgggggcat gaatgactga 2100
gccacacatt cagaggtggg ttgtgaagtc ttccactgtg aactatgggt tttttttttt 2160
tttttttttt ttgggtgtgt tgtgtaaagt tgctaagcca tgggggtggga ggggaaagat 2220
ttactagggtg caaagtgggc aagggccaga tatactataa ttgtctatat tatagagtta 2280
gtatacatct gtattttttc ttgacacact tttgcttggt ggtgctttcg ttaaaattac 2340
actttaattg cttccg 2356

<210> 22

<211> 3411

<212> DNA

<213> Homo sapiens

<400> 22

aagatactta gggaaagcga agaagagcat gctgctacac accactctgg cttacgcttc	60
cagctaacag cggtttgtct gtgtttctgg actgaagtga ctgatgaaaa acagtagtca	120
agtctggagg gtgattccct ttcagttcac ttgcctttca gttctgggga aattagaaga	180
tttgcctctt ctctctgctt cttagaatgg gctgcaatth gtgcactttt cagaaaagag	240
aggaacacta caaactgctg tatgaagttt cccaggagct gtgacttgaa caatgctatt	300
ctgcatgaaa ttgagaaatt gccagtgttc ccaacctgat ggagagtatg tgtgggatct	360
tagttctgct tcattaagct ctttgtgaat gagtggtaag agacaatttt tccttaaacc	420
tttggcattg ttgacatgtt ggccatccac tgtcaatggg aaggatcttt caaaggccac	480
tcataagag gcagtggag cttttcgcaa tgccaaggag cccattgtgg tgcaggtgtt	540
aaggcgaaca cctcttagta gaccagccta tgggatggct tcagaagtgc agcttatgaa	600
tgccagcact cagacggaca tcacctcga acacatcatg gctctggcca agcttcgtcc	660
acctaccct ccagtgccag acatctgtcc attcctgtc tcagacagct gccattctct	720
acatccaatg gagcatgaat tttatgagga caatgagtat attccagct tgcctgtga	780
tgcagacaga acagaagact ttgaatatga ggaggtcgag ttgtgtcgtg ttagcagtca	840
agagaagctg ggcctgacag tctgttaccg aacagatgat gaagaagaca ccggcattta	900
tgtcagcgag gttgaccaa atagcattgc tgccaaagac ggccggattc gagaagggga	960
tcggattttg caaataaatg gggaagatgt ccagaatcga gaagaagcag tggccttgct	1020
gtctaacgat gagtgtaga gaatcgtgct gcttgttgca aggccagaga ttcagctgga	1080
tgaaggctgg ctggaagatg aaaggaatga attcttagag gagttaaact tggagatgtt	1140
ggaagaagag cataatgaag caatgcagcc cactgccaat gaggtggagc agccaaaaaa	1200
gcaagaagaa gaagaaggca caacagacac tgcaacatcc tcatccaaca accatgagaa	1260
ggacagtgga gtaggacgta cagatgaaag cttgcgaaat gatgagagct cagagcagga	1320
gaatgcagcc gaggaccca atagcacatc tttgaagagc aagagagacc tggggcagag	1380
ccaagacact ctgggaagtg ttgaacttca gtacaatgag agcctcgtat ctggtgaata	1440
cattgactca gactgcattg gcaaccaga tgaggactgt gaaagattca ggcagctctt	1500
ggagctcaaa tgcaagattc gaaatcatgg agagtatgac ctgtattact caagcggcac	1560
aattgaatgc aatcaagggg agcaagaggg agtggagcat gagctacagt tgcttaatga	1620
agaactgaga aacattgagc ttgagtgtca gaatatcatg caggctcaca ggctccagaa	1680

agtgacagac cagtatggag acatctggac attgcatgat ggaggattcc ggaattataa 1740
caccagcata gatatgcaaa ggggaaagct agatgacatc atggagcatc cagaaaagtc 1800
tgacaaggac agttctagtg cttacaacac agctgagagc tgcagaagta ctccgctcac 1860
tgtagaccgt tcccctgaca gttcccttcc aagggtgatc aacctcacca ataagaaaaa 1920
cctgagaagc acaatggcag ccaccagtc ctcttccgga cagagcagta aagagtcgac 1980
ctccacaaaa gccaaaacca ctgagcaagg ttgtagcgct gaaagcaagg agaaggtttt 2040
agaaggcagc aagcttcttg atcaagagaa ggcagtcagc gaacacatcc cttacctctc 2100
tccttaccac agtctctcat atagatatgc aaacatccca gcacacgccc ggcattatca 2160
aagctacatg cagttaattc aacagaaatc tgcagtcgag tatgctcaga gtcagctcag 2220
cttggtgagc atgtgcaagg agtctcagaa gtgttcagag cccaagatgg aatggaaggt 2280
gaaaattagg agcgacggga cacggtacat caciaagaga cccgtgagcag accggatcct 2340
gaaggaacgt gccttaaaga tcaaggaaga gcggagtggc atgaccacag acgatgacac 2400
catgagcgag atgaaaatgg ggcgctactg gagcaaagag gagagaaagc agcacctggt 2460
tagggccaaa gagcagcgcc gtcgccgtga gttcatgatg cgaagcaggt tagagtgtct 2520
caaggagagc cctcagagcg gcagtgaggg caagaaggag atcaatatca ttgaactgag 2580
tcacaaaaag atgatgaaaa agagaaacaa gaaaattttg gacaactgga tgacaatcca 2640
agaactgatg acccatgggg ccaagtctcc agatggcacg agagtccata atgccttctt 2700
gtcggtgacc actgtatgac cgaatgaatg gaatgcatgc gactgatttt aggaggatgc 2760
taccagtttc ggtagagtat gattgcctcg ttcaatgtgg cgtttttata tatattttgt 2820
gactctttat agtttaaatt ttttgtaagc aaaaaatacc tggtaatttt tcattttgtt 2880
ttcatatact ggtaccttct ttttggtgga gatctttctt ttacttgtga tatattcata 2940
ttactcgttt ataaaaaatc aaaaacaaaa ggaaagaaaa caaaaaaac ttgcacaaaa 3000
atactgagga gccaatat ttcctacttc aatgtatcta atgtaagtga aaatctggat 3060
ttatttctct agtttactta ttttctactt taataataac tcaatgccaa aatatttcta 3120
tgcttgtttc ttggcataat acatttttaa tcaaacctgt tcgtaataaa tcatgtgcca 3180
catcttgtct tacacataaa agccactgct tttaacatcc cattgtacat ttaacacata 3240
aatgggtgta acatcccat gtacatttaa cacataaatg gccattgtct ttgtctcagt 3300
aaagtgtgag tggacaatct tcctgtgata tgtagtgaca ttggtcatgt agctagataa 3360
ttatggtaat tgtgacaatt aaagagaaat aaattattga ttatccttca g 3411

<210> 23

<211> 2078

<212> DNA

<213> Homo sapiens

<400> 23

actatattga atagaagcga agagacagag tatacttgtc ttgttcctaa tccagaagaa	60
aagtttgaat cttttacat tgagtaagct tgagtgccca agggtttgat aaatgtggaa	120
aataaccgct gttcgaatat cctgttggcc tgctgaagaa atgactggct gaggctgtag	180
agaaatcttg tggtcattct catggaaaga ggattggagc cccgattctt tgatctcatt	240
attgaccagt tgattgagaa gactgtcctc ctgtccccac ccaatcaca taagtggctc	300
tgaaatcaca tgaagctgtg gatggagagt cacctgatag tcccagaaac ccgtcccagc	360
ccaaggatga tgagtaacca gacgttggta accgagttca tcctgcaggg cttttcggag	420
caccagaat accgggtgtt cttattcagc tgtttcctct tcctctactc tggggccctc	480
acaggtaatg tcctcatcac ctgggccatc acgttcaacc ctgggctcca cgctcctatg	540
tactttttct tactcaactt ggctactatg gacattatct gcacctctc catcatgccc	600
aaggcgctgg ccagtctggt gtcggaagag agctccatct cctacggggg ctgcatggcc	660
cagctctatt tcctcacgtg ggctgcatcc tcagagctgc tgctcctcac ggtcatggcc	720
tatgaccggt acgcagccat ctgccacccg ctgcattaca gcagcatgat gagcaagggtg	780
ttctgcagcg ggctggccac agccgtgtgg ctgctctgcg ccgtcaacac ggccatccac	840
acggggctga tgctgcgctt ggatttctgt ggccccaatg tcattatcca tttcttctgc	900
ggggtccctc ccctgctgct tctctcctgc agctccacct acgtcaacgg tgtcatgatt	960
gtcctggcgg atgctttcta cggcatagtg aacttcctga tgaccatcgc gtcctatggc	1020
ttcatcgtct ccagcatcct gaaggtgaag actgcctggg ggaggcagaa agccttctcc	1080
acctgtcttt ccacctcac cgtggtgtgc atgtattaca ccgctgtctt ctacgcctac	1140
ataagcccgg tctctggcta cagcgcaggg aagagcaagt tggctggcct gctgtacact	1200
gtgctgagtc ctacctcaa cccctcatc tatactttga gaaacaagga ggtcaaagca	1260

gccctcagga agcttttccc tttcttcaga aattaacttg tgtcttctga agtttttgtc 1320
 cttggagact gcagtttttag tggagtctga tgagcaagtt tctggactgg gaggtaaata 1380
 gacttgtttt ttctaacca gcagaacctg tgcattgatg agctgggagt gctgggtccc 1440
 tgagacttgg tgatgggggg attgcggtag gtgtgacca ccagcctctg agacaccaga 1500
 tctcttggtc tgactccaag gtggtgtcag tgctgagttg gcagagaaac acagggccta 1560
 ggaggagaaa agtgttcctt gtgttttcaa ctcatagcca catcaaact tttccttttg 1620
 tatgtacaac actcagggtg taaattctaa aatcacacct agaattcagc tgtgtgggtc 1680
 taaaaagtta ggcagaattg tggacccac agccctgaa gggttgtcag actgaagctt 1740
 tgttttaagg tgaccactgg cctctaggaa accttctacc ccattatct tacctcaact 1800
 gtttcctctg gagccacttg gacttcaaaa atagtttgat tttaatgcaa actgttgttt 1860
 ctgcttttgt cttttcttct tactaacact tcagatcctg agtatctgac ctacctaat 1920
 ggcctggagc gaggacagct gccatgagtc agctctaata agctaaacgg cattgtagta 1980
 aggcatgtct atccttccat tgtttcagat aaaactgggt gaacgaatcc tctgtgtcta 2040
 ataaagacag gtcttgtgct tgtgtccat gctattac 2078

<210> 24

<211> 723

<212> DNA

<213> Homo sapiens

<400> 24

agaccagacc aacagtaaca ccaagggcag gtgggcaggc ctccgcctc ctcccctact 60
 ccagggccca ctgcagcctc agcccaggag ccaccagatc tccaacacc atgggtccgat 120
 accgcgtgag gaggctgagc gaacgctcgc acgagggtga caggcagcag ttgcatgggc 180
 aagagcaagg acaccacggc caagaggagc aagggtgag cccggagcac gtcgaggtct 240
 acgagaggac ccatggccag tctcactata ggccgagaca ctgctctcga aggaggctgc 300
 accgatcca caggcggcag catcgctcct gcagaaggcg caaaagacgc tcctgcaggc 360
 accggaggag gcatcgcaga gagtccctag gtgacccct caaccagaac tttctttccc 420

aaaaggctgc agaaccagga agagaacatg cagaaggcac taagcttcct gggccctca 480
 cccccagctg gaaattaaga aaaagtcgcc cgaaacacca agtgaggcca tagcaattcc 540
 cctacatcaa atgctcaagc ccccagctgg aagttaagag aaagtcacct gcccaagaaa 600
 caccgagtga ggccatagca actcccctac atcaaatgct caagccctga gttgccgccg 660
 agaagcccac aagatctgag tgaaatgagc aaaagtcacc tgcccaataa agcttgacaa 720
 gac 723

<210> 25

<211> 4352

<212> DNA

<213> Homo sapiens

<400> 25

gatctgcgca tgcgcactcg aggcaaggcg gccactccca gaagtctccg gggctcgagc 60
 ggtgggtggg aggcgcggtc caagtggctg ctgccgtatt tctcgtaggt ttttaggatt 120
 ctggactcca tttctcaccg gtacctgcgg tcctgactgt tctggcttgg actacatatt 180
 ccaaaggcaa cggagaaaac ttggaccctt ggctccggtt ccgggacctt ttggttccac 240
 atgaagaagg cggcgcaggt cactgcggga gcctctgcag cccagcccg ccgcccgcgt 300
 ggtaccctgc gtggcgggtga gagtggggct gcaggaggcg ctcagagcgc ggaggcccc 360
 aggcgggccc ggcgagtctc ccgcctcccc aggccctagg gtgctggagc tgcggaggag 420
 gggcccgtct ctaggagaca ggcggggagc gtgagagacc gcgcagcggc gacaggaccg 480
 gcttgctgt gtgacgtggg tgcagttgcg ggactgtgag gggctcttgg tcttttgact 540
 ttgctcttct ggtagagaaa cgccaaggac tgatcaactc ttgccagtct caaacatgg 600
 cctatggatc aataacattc ggggatgtgg ccatagactt ctctcatcag gagtgggaat 660
 atctcagcct gggttcagaag accttgtagc aggaggtgat gatggagaac tatgacaact 720
 tagtctcatt ggcaggacat tccgtatcta agccagattt aatcacctta ttggagcaag 780
 gaaaagagcc atggatgatt gtgagggaag aaacaagagg agagtgtaca gatttgatt 840
 caagatgtga aataatcagc gatgggaaaa tgcagcttta taggaaacac tcatgtgtta 900

ctctccatca gagaattcat aatggacaga aaccatatga gtgtaagcaa tgtcagaagt 960
ccttttagtca tcttacagaa ctcatggtag atcaacaat tcatactagt gaggaacctg 1020
atcaatgtga aaagtttagg aaggcattta gccatcttac agaccttaga aagcatcaga 1080
aaattaatgc tcgtgagaaa ccttatgaat gtgaagaatg tgggaagggtc ttcagttatc 1140
ctgcaaacct tgctcaacat gggaaagttc atgttgagaa accctatgaa tgtaaagagt 1200
gtggggaagc ttttaggact agccgtcaac ttactgtaca tcatagattt cattatgggtg 1260
agaaacccta cgaatgtaag gaatgtggca aggccttttag tgtgtatgga cgacttagtc 1320
gacatcagag tattcacact ggtgagaagc cctttgaatg taacaaatgt gggaagtcct 1380
ttaggtcaa agcaggcctt aaagtacatc agagtattca tactggagag aagccacatg 1440
aatgtaagga atgtggaaag gcctttcgtc agttttccca ccttgtgggt cataaaagaa 1500
ttcatactgg agaaaaacc tatgaatgca aggaatgcgg gaagggtttt acatgtaggt 1560
atcaacttac catgcatcag agaatttatt caggggagaa acactatgaa tgtaaagaaa 1620
atggggaggc ttttagtagt ggccatcaac ttactgcacc tcatacattt gaaagtgttg 1680
agaaacctta taagtgtgag gaatgtggga aagccttttag tgtgcatgga cgacttactc 1740
gacatcaggg tattcatagt ggtaagaaac cctatgaatg taacaaatgt gggaagtcct 1800
ttaggtcaa ttcacccctt aaaatacatc aaaatattca taccggtgag aaacctaca 1860
aatgtaagga atgtgggaag gccttcagtc agcgtgcaca ccttgcccat cataacagaa 1920
ttcatactgg ttacaaacc tttgaatgta aagaatgtgg gaagtccttt cgttgtgcct 1980
catatcttgt tatacatgag agaattcata cgggagagaa accctatgta tgtcaagagt 2040
gtgggaaggg ttttagttat agccataaac tcactataca tcgcagagtt catactgggtg 2100
agaaacctta tgaatgtaag gaatgtggga aggccttttag tgtatctgga caacttactc 2160
agcatctgag tattcacagt ggtaagagac cctttgaatg caacaaatgc gggaagtcct 2220
ttaggttcat ttctgtactt aaagcccatc agaattattca tagcgctgag aaacctacg 2280
aatgtaagga atgtggcaag gcctttcgtc atgccacaag cctcatatat catgaccgaa 2340
ctcatgctgg tgaaaagtcc tatgaatgta aagaatgtgg ggaaactttt agtcatgctt 2400
cacatcttat tattcatgag agaattcata ccagtgataa accctatgaa tgtaaagat 2460
gtgggaaggc atttcactgt gcctcatatc ttgttagaca tgaaagcgtt catgctgatg 2520
gaaatcccta tatgtgtgaa gagtgtggga aagcttttaa tagtagccat gaacttagta 2580
tacatcatag agttcatact ggtgagaaac cctttaaatg taacaaatgc agaagtcct 2640

ttaggcttag atccatcctt gaagtacatc agagaattca tatttgatag agaaattgtg 2700
 gcatatacag gtagtaaaga aagcacattg tgtcaatgga tttatccac tgaacaattt 2760
 gaatgaagac aatgtggaat ggctttgaga tacagttcag tttttattag gcatcaattt 2820
 gtattggtgt gaacttttca aaatgtaa atcatatgca gtttctttgg tatgacctgt 2880
 gttaactcag attcagagaa tttataatgt atttttttaa taacataaga aaacttttct 2940
 gatattcctc tagttttgga ttgccaaaat attcatagta gaggaaaact gagaaaacgt 3000
 aggaaacat tctttatact tgcagtataa catcttccag caacaatcat ttattgttaa 3060
 acctggaaaa agccattaca caggaaacct ctgttgatat aagcgatgat gggaaaggct 3120
 atttagccac tctgcatatt taatgctcat gatcaacatc tgcttcctat agctgggatg 3180
 atatggggga agatatttat cccttctggg gatcatatat aaatgatgat agtagtacct 3240
 acagtatagt gctgttagaa ttacatgagt tagatgtgga ggtcagagtg gaagcaggtg 3300
 tgagagggtc ccgcagaaga aaacatggct gccaaagtgt ttgagtccat cggcaagttt 3360
 ggcttggcct tagctgttgc aggaggcatg gtgaactctg ccttatgtaa tgtggatgct 3420
 gggcacagag ctgccatctt tgaccaattc cgtggagtac agaacattgt ggtaggggaa 3480
 gggactcact ttctcatccc atgtgtacaa aaaccaatta tctttgactg ctgttctcaa 3540
 ccacgtagtg cgccagtcac cactggtagc aaagatttac agaattgtcaa catcacactg 3600
 tgcacacctt tccggcccat cactagccag ctctctcgca tcttcaccag cattggagag 3660
 gactacgatg agtgtgtgct gccgttcatt accacggaga tcctcaagtc actggtggct 3720
 cgctttgatg ctggagaact aatcaccag agggagctgg tctccagcca ggtgagcaac 3780
 aaccttatgg agtgagcagc cacctttggg ctcatcttgg acgacgtgtc tttgacacat 3840
 ctgaccttct tgaaggagt taaacagacagg tggccccca ggaagcagag agcgccagat 3900
 ttgtggtgga aaaggcggc atcatctctg ctgagggtga ctccaaggca gctgagctga 3960
 tcgccaactc actggccact gcaggggacg gccagagcga gctgtgcaag ctggaagctg 4020
 cagaagacat tgcataccag ctctcatgct ctcggaacat cacctgcctg ccggcagggc 4080
 agtccgtgct cctccagctg ccctagttag gccccagcct acctgcacct ccgtgaggca 4140
 actgggccac agccccgatg attcgtaaca ccacctttcg ccctcaccac agaaatcact 4200
 gaaatttcat gattggctta aagtgaagga agtaaaggta aaatcacttc agatctctaa 4260
 aaaacaaaga attacatgag ttagtacatg aaaaaattat gggaagctac atgaaatata 4320
 ctgttacggt caataaacat tagcttctgt at 4352

<210> 26

<211> 3370

<212> DNA

<213> Homo sapiens

<400> 26

cttctggaac	cgagctgtgg	cctggggttc	catggagcag	gtggggcgcg	gaggaactga	60
gctgtggcct	ggggttccgt	ggagcaggtt	gagtatggag	gaactggggg	tccgtggagc	120
aggtggggcg	cagaggaacc	gagctgtgga	atggggttcc	gtggagcagg	ttgagtgtgc	180
cggctcgtgga	cgggtcgtgg	gcgggactcc	ccagtgtgtc	accactgagg	catcttcgcc	240
ctcttacgtg	tctctgtttc	tctctctaga	agtccaacac	ctccgagagg	actcaaaaca	300
cagagtgacc	agtcctatg	tgcatttcct	gggcgacagc	gagactccat	ctcaataaat	360
aaaaaaaaaa	aagagttgtt	atcagtagaa	gggaatgtct	ggttacagta	tggcgttgtg	420
cagatgaagg	tcttatcgca	gatgaagcca	ccaggtcaca	agcctcagag	agaatcaact	480
ataaatgctt	ctcatcagac	tcaaggcctg	aggtgatgct	gatgctgtgc	ctgaattcca	540
gcaggggagga	ggcatgaggg	cagtgtgtgc	acagaagaca	acaccgtcc	ctcgttacct	600
gtggccccggc	cacctcagcg	gccaaggag	gtcacctgg	tcatggtgca	gtgaccacag	660
gacccccaca	tgccgggagc	tgggttcgcc	ccaccccacc	ccctgcaccg	ggccagcgag	720
gggatggccc	agaagagggg	gaggaccatg	tggattcacc	agtgtggac	atgtgctctg	780
tggctacccc	ctctgcctac	tctctggccc	gatacagggg	tgtgggacag	gcctgggtga	840
ctccagcatg	gctttcctct	ccaggacgtc	accggtggca	gctgcttcct	tccagagccg	900
gcaggaggcc	agaggctcca	tcctgcttca	gagctgccag	ctgccccgc	aatggctgag	960
caccgaagca	tggacgggag	aatggaagca	gccacacggg	ggggctctca	cctccagatc	1020
gcctgggcct	gtggctcccc	agaggccctg	ccacctgaag	ggatggcagc	acagacccac	1080
tcagcacaac	gctgcctgca	aacagggccca	ggctgcagcc	cagacgcccc	ccaacacggg	1140
ccaccatcag	caccaccacc	accacccaag	gaggggaccc	aggaggggct	ggtggagctg	1200
ccgcctcgt	tccgggagct	gtcaccttc	ttctgcacca	atgccaccat	ccacggcgcc	1260

atccgcctgg tctgctcccg cggaaccgc ctcaagacga cgtcctgggg gctgctgtcc 1320
ctgggagccc tggtcgcgct ctgctggcag ctggggctcc tctttgagcg tcaactggcac 1380
cgccccgtcc tcatggccgt ctctgtgcac tcggagcgca agctgctccc gctggtcacc 1440
ctgtgtgacg ggaaccacg tcggccgagt ccggtcctcc gccatctgga gctgctggac 1500
gagtttgcca gggagaacat tgactccctg tacaacgtca acctcagcaa aggcagagcc 1560
gccctctccg ccaactgtccc ccgccacgag ccccccttcc acctggaccg ggagatccgt 1620
ctgcagaggc tgagccactc gggcagccgg gtcagagtgg ggttcagact gtgcaacagc 1680
acgggcggcg actgctttta ccgaggtac agtcaggcg tggcggctgt ccaggactgg 1740
taccacttcc actatgtgga tctctggcc ctgctgcccg cggcatggga ggacagccac 1800
gggagccagg acggccactt cgtcctctcc tgcagttacg atggcctgga ctgccaggcc 1860
cgacagttcc ggaccttcca ccacccacc tacggcagct gctacacggt cgatggcgctc 1920
tggaacagtc agcgcgcccg catcacccac ggagtcggcc tggctctcag ggttgagcag 1980
cagcctcacc tccctctgct gtccacgtg gccggcatca gggcatggt tcacggccgt 2040
aaccacacgc ctttcttggg gcaccacagc ttcagcgtcc ggccaggac ggaggccacc 2100
atcagcatcc gagaggacga ggtgcaccgg ctcgggagcc cctacggcca ctgcaccgcc 2160
ggcggggaag gcgtggaggt ggagctgcta cacaacacct cctacaccag gcaggcctgc 2220
ctggtgtcct gcttccagca gctgatggtg gagacctgct cctgtggcta ctacctcac 2280
cctctgccgg cgggggctga gtactgcagc tctgcccggc accctgcctg gggacactgc 2340
ttctaccgcc tctaccagga cctggagacc caccggctcc cctgtacctc ccgtgcccc 2400
aggccctgca gggagtctgc attcaagctc tccactggga cctccaggtg gccttccgcc 2460
aagtcagctg gatggactct ggccacgcta ggtgaacagg ggctgccgca tcagagccac 2520
agacagagga gcagcctggc caaaatcaac atcgtctacc aggagctcaa ctaccgtca 2580
gtggaggagg cccccgtgta ctcggtgccg cagctgctct cggccatggg cagcctctgc 2640
agcctgtggt ttggggcctc cgctctctcc ctcctggagc tcctggagct gctgctcgat 2700
gcttctgccc tcacctggt gctaggcggc cgccggctcc gcagggcgtg gttctcttgg 2760
cccagagcca gccctgcctc aggggcgtcc agcatcaagc cagaggccag tcagatgccc 2820
ccgcctgcag gcggcacgtc agatgaccgc gagcccagcg ggcctcatct cccacgggtg 2880
atgcttccag gggttctggc gggagtctca gccgaagagg gccgggctgg gccccagccc 2940
cttgagactc tggacacctg aaccagacct gccagggtg tgcgatctct tggcctggtc 3000

cttgcagctg tggcagcagc aggctcccca gcggcccagg gtgggccaga ccagcagccc 3060
 aggaagcagc acacgcggcc gtggggaggc aggcaccggg catgtcggcg cctctggtca 3120
 aaccacctac actgcctggg gtgggtctca aggaggcccg gggcggaggg gggttcccg 3180
 gtgcacacga gtgcggctgg acgtgccgac acgcggtgat gtacccatgc tccgtgtgtc 3240
 tgtgtctgca tgtccacacg tctgatgcac ctgtgtacgt gtgtcaagcc tagccacctc 3300
 agctgcaggg aggcagaagg caaggcaggc cccacggaca cacttgggct gctctgaaat 3360
 aaagctgttg 3370

<210> 27

<211> 4003

<212> DNA

<213> Homo sapiens

<400> 27

attgaccctt ggctccaggt tccctgctgc ttcctgtgcc ttcctcagcc tagtacgact 60
 acaacagaga aatggtccca gtgggaaatc agatactgag tcttagatga agatttccag 120
 tgagaagaaa cagataggat gagaagctta gaagagtaat tgcccttaag aaaggtcttg 180
 gttcttttga ctttgcctct ctggtagaga aacgccaagg actgatcaac tcttgccagt 240
 ctcaaaccat ggcctatgga tcaataacat tcggggatgt ggccatagac ttctctcatc 300
 aggagtggga atatctcagc ctggttcaga agaccttgta ccaggaggtg atgatggaga 360
 actatgacaa cttagtctca ttgggacatt ccgtatctaa gccagattta atcaccttat 420
 tggagcaagg aaaagagcca tggatgattg tgagggaaga aacaagagga gagtgtacag 480
 atttggattc aagatgtgaa ataatcagcg atgggaaaat gcagctttat aggaaacact 540
 catgtgttac tctccatcag agaattcata atggacagaa accatatgag tgtaagcaat 600
 gtcagaagtc ctttagtcat cttacagaac tcatggtaca tcaaacaatt catactagt 660
 aggaacctga tcaatgtgaa aagtttagga aggcatttag ccatcttaca gaccttagaa 720
 agcatcagaa aattaatgct cgtgagaaac cttatgaatg tgaagaatgt gggaaggtct 780
 tcagttatcc tgcaaaccct gctcaacatg ggaaagttca tgttgagaaa ccctatgaat 840

gtaaagagtg tggggaagct tttaggacta gccgtcaact tactgtacat' catagatttc 900
attatggtga gaaaccctac gaatgtaagg aatgtggcaa ggccttttagt gtgtatggac 960
gacttagtcg acatcagagt attcacactg gtgagaagcc ctttgaatgt aacaaatgtg 1020
ggaagtcctt taggtcctaa gcaggcctta aagtacatca gagtattcat actggagaga 1080
agccacatga atgtaaggaa tgtggaaagg cctttcgtca gttttccac cttgtgggtc 1140
ataaaagaat tcatactgga gaaaaaccct atgaatgcaa ggaatgcggg aagggtctta 1200
catgtaggta tcaacttacc atgcatcaga gaatttattc aggggagaaa cactatgaat 1260
gtaaagaaaa tggggagggt tttagtagtg gccatcaact tactgcacct catacatttg 1320
aaagtgttga gaaaccttat aagtgtgagg aatgtgggaa agccttttagt gtgcatggac 1380
gacttactcg acatcagggt attcatagt gtaagaaacc ctatgaatgt aacaaatgtg 1440
ggaagtcctt taggtcctaa tcattccctta aaatacatca aaatattcat accggtgaga 1500
aacctacaa atgtaaggaa tgtgggaagg ccttcagtca gcgtgcacac cttgcccatc 1560
ataacagaat tcatactggt tacaaccct ttgaatgtaa agaattgtggg aagtcctttc 1620
gttgtgcctc atatcttgtt atacatgaga gaattcatac gggagagaaa ccctatgtat 1680
gtcaagagtg tgggaagggt tttagttata gccataaact cactatacat cgcagagttc 1740
atactggtga gaaaccttat gaatgtaagg aatgtgggaa ggccttttagt gtatctggac 1800
aacttactca gcatctgagt attcacagt gtaagagacc ctttgaatgc aacaaatgcg 1860
ggaagtcttt taggttcatt tctgtactta aagcccatca gaatattcat agcgctgaga 1920
aacctacga atgtaaggaa tgtggcaagg cctttcgtca tgccacaagc ctcatatc 1980
atgaccgaac tcattgctggt gaaaagtcct atgaatgtaa agaattgtggg gaaactttta 2040
gtcatgcttc acatcttatt attcatgaga gaattcatac cagtataaaa ccctatgaat 2100
gtaaaagatg tgggaaggca tttcactgtg cctcatatct tgtagacat gaaagcgttc 2160
atgctgatgg aaatccctat atgtgtgaag agtgtgggaa agcttttaat agtagccatg 2220
aacttagtat acatcataga gttcactg gtgagaaacc ctttaaagt aacaaatgca 2280
gaaggctcctt taggcttaga tccatccttg aagtacatca gagaattcat atttgataga 2340
gaaattgtgg catatacagg tagtaaagaa agcacattgt gtcaatggat ttatcccact 2400
gaacaatttg aatgaagaca atgtggaatg gctttgagat acagttcagt ttttattagg 2460
catcaatttg tattggtgtg aacttttcaa aatgtaaacc atatatgcag tttctttggt 2520
atgacctgtg ttaactcaga ttcagagaat ttataatgta tttttttaat aacataagaa 2580

aacttttctg atattcctct agttttggat tgccaaaata ttcatagtag aggaaaactg 2640
agaaaacgta ggaaaccatt ctttatactt gcatgtaaac atcttccagc aacaatcatt 2700
tattgttaaa cctggaaaaa gccattacac aggaaacctc tgttgatata agcgatgatg 2760
ggaaaggcta tttagccact ctgcatatct aatgctcatg atcaacatct gcttcctata 2820
gctgggatga tatgggggaa gatattttatc ccttctgggg atcatatata aatgatgata 2880
gtagtaccta cagtatagtg ctgttagaat tacatgagtt agatgtggag gtcagagtgg 2940
aagcaggtgt gagagggtcc cgcagaagaa aacatggctg ccaaagtgtt tgagtccatc 3000
ggcaagtittg gcctggcctt agctgttgca ggaggcatgg tgacctctgc cttatgtaat 3060
gtggatgctg ggcacagagc tgccatcttt gaccaattcc gtggagtaca gaacattgtg 3120
gtaggggaag ggactcactt tctcatccca tgtgtacaaa aaccaattat ctttgactgc 3180
tgttctcaac cacgtagtgc gccagtcac actggtagca aagatttaca gaatgtcaac 3240
atcacactgt gcatcctctt ccggcccatc actagccagc ttcctcgcat cttcaccagc 3300
attggagagg actacgatga gtgtgtgctg ccgttcatta ccacggagat cctcaagtca 3360
ctggtggctc gctttgatgc tggagaacta atcaccaga gggagctggt ctccagccag 3420
gtgagcaaca accttatgga gtgagcagcc acctttgggc tcattctgga cgacgtgtct 3480
ttgacacatc tgaccttctt gaaggagtta acagacaggt ggccccccag gaagcagaga 3540
gcgccagatt tgtggtggaa aaggcggcca tcatctctgc tgagggtgac tccaaggcag 3600
ctgagctgat cgccaactca ctggccactg caggggacgg ccagagcgag ctgtgcaagc 3660
tggaagctgc agaagacatt gcataccagc tctcatgctc tcggaacatc acctgcctgc 3720
cggcagggca gtccgtgctc ctccagctgc cctagtgagg ccccagccta cctgcacctc 3780
cgtgaggcaa ctgggccaca gccccgatga ttcgtaacac cacctttcgc cctcacccca 3840
gaaatcactg aaatttcatg attggcttaa agtgaaggaa gtaaaggtaa aatcacttta 3900
gatctctaaa aaacaaagaa ttacatgagt tagtacatga aaaaattatg ggaaactaca 3960
tgaaatatac tgttacgttc aataaacatt agcttctgta tat 4003

<210> 28

<211> 4199

<212> DNA

<213> Homo sapiens

<400> 28

ccttgagtca	gccatggggt	cagccacaca	gtcaccgcag	ctgctctttg	gcaccgggcg	60
ctggaaagag	cctaggatga	cacagcctgg	aaagagcttg	ggagaagctc	atcttccaca	120
gaactacctg	ctataccagc	cagggcaggt	gcttattccc	cacaacagcc	ctctgttgta	180
ggcggcagtg	ccatcctgaa	ggtgccgtgg	taccttctga	agacgcaggc	tgaggcctgt	240
aatggcactg	ccatgcacac	atggcacacc	ctctcccggg	taaactgctt	gcccttcaact	300
gggccctggg	tgcccaggga	ggtgacgcta	gggtgcctgg	ctgcgtctct	tccacaccca	360
cggagaacgg	agttgtgttc	tgggggtccag	gctgccacag	cctgctggta	taccagaggg	420
tgcatacgag	ggagtgggca	gggccaagag	ctttcaaatg	ctccagtggg	cttcatattt	480
tgcatttggg	ggagtgggtc	agacaatggc	tgtattgctg	ctgggtgggt	ggacaggttg	540
ccctgaaagg	aggcaagcag	cagagcctga	cagtcagatg	tgggcccact	gagcagccat	600
gtggccctgg	acagagcaga	tacttgggcc	tcagtgtcct	caagctgtgg	agcacagtta	660
gtgaaataaa	agtgtttctt	cccatgctgt	cctaaagcgg	gtgctccgtt	agatgcctgc	720
tccttttcgc	tttacagatg	aggaaactct	cacagagagg	tgaattaagt	ggccatcctt	780
catttagcaa	atacttgggt	ccgtgctggg	cactggggct	cagcagtga	caccgcaggg	840
aggcacctgc	acctgagaac	ttcccttcta	gaggggagac	ggggagtcac	aagtagacaa	900
aggagggggac	ctgcaagtag	tcataccggc	tgtgaaggag	gcaggtgggg	tgtggggacg	960
gtgacaggag	gagcggccag	ctcatctgga	gccagctagt	tatctgctgt	gccagtgttc	1020
ttgacccccct	gcttgtgagc	tgctgtgctg	ggtgctgggg	ctgcagacc	atgcctggct	1080
cccaggggct	cgccatctgg	ccgggcgggg	caggtgctgt	gagcaggtgc	tgagtgagcc	1140
ctctccttgg	ggcaggagag	gtgcatcctg	aggaaacagg	gagcggggag	cacgggcagg	1200
agctcagagg	acgactctgg	ggggactgga	agtgaagggtg	gcaggtggag	ggtcccaggt	1260
gggattccag	gctgggtgagc	cacggtgggtg	ctgtggcctg	aagacagtgt	ggtatgtgtg	1320
tgcccgtagg	ccggccgtac	cgtaggcaagc	ccagtgacat	gtgggccctg	ggcgtggtgc	1380
tcttcacat	gctgtatggc	cagttcccct	tctacgacag	catcccgcag	gagctcttcc	1440
gcaagatcaa	ggctgccgag	tataccattc	ctgagtgagt	cctgcctccc	gggggggtggg	1500
gggggcagcc	agcctgatgg	gcttaaggca	gggaaacggc	acagtgagcc	tgggcatacct	1560

ggccacagac ccagcaccac ccactgagcc ctctctccca actccctgcc tcagggatgg 1620
acgggtttct gagaacaccg tgtgtctcat ccggaagctg ctggtccttg acccccagca 1680
gcgccctggcc gccgccgacg tcctggaggc cctcagtgcc atcattgcat catggcagtc 1740
cctgtcatct ctgagtgggc ctttgcaagt ggttcctgac attgatgacc aaatgagcaa 1800
tgcggatagc tcccaggagg cgaaggtagc ggaggagtgc tcccagtagc agtttgagaa 1860
ctacatgcgt cagcagctgc tgctggccga ggagaagagc tccatccatg acgcccggag 1920
ctgggtaccc aagcggcagt tcggcagcgc accaccggtg cgacggctgg gccacgacgc 1980
acagcccatg acctccttgg acacggccat cctggcgcag cgctacctgc ggaaataaca 2040
gcctcagccg gggccaccag cactgctgcc acttcttcca gcccagcca aaggcgtggc 2100
tgtcagggtc gggccctgta gtgctggact ctcccgggcc acaataggga cagggcaggg 2160
acagggacag cccaggtcac acgtggggtc agcagaggta ccacgaagct accttttggg 2220
atgattgctc gattgtttgg tttttaaatc tgagaagcct agataactaa tctgctttta 2280
atcacgatgt tttaatctac ctctgtctct ttaaccatgc tgtctctgga ctgagcaaga 2340
gggaggaggg agcctgctca cccactcca gggccttccc cagcggccac caactgacct 2400
ggggcgctgc tccccacagt ccaaataagc tgaaagtga gctcgtgca ggccccagag 2460
cgagcttccc ctctccctg ctctcccagg cccctgccac agcctctttc cgtccctctc 2520
tttctgatcc aggccctca gtccaagctt tggaacacct tcacctcatc ttaaaccgaa 2580
ctcaaataata tttatttttt taccatacca acttctctcc catctctagg tggctcagtc 2640
catggccact ccctgcccc agcctggctg gacagcaagg aatccacagc ccacacgtga 2700
gtccctcct cacccccagg cagggaagcc cctcctgcc gtcctgtcc cttttcagcc 2760
caccagtccc tctctgctgc cggatgatggg aggcccttct agacctggct ctttctctcc 2820
cgtctcagtg gcttctctga ggtgctgtac acgcgcgta acctgttccc ttctctatcc 2880
ttccccgtgg tactgagctc acgtggactc ccagtgcgaa ggggcccatt ggttgggctg 2940
caggcctggc cgtgagcggg ggctgcctgc acgtccctc agcctactct tgtgtttagg 3000
ggatggtggg aacatatccc agtgcccttg cctcataata gatgtggtga ctctcccgtt 3060
agaccctagc aagggtcctc catggtgggtg agggactcag gagaattgta gggattgggg 3120
gaccctgcct gcctggcttg agaacagccc tgctgccctt ttgagccgag attttgaagt 3180
ggatgcccgt cttgccagaa atgctgttct caccagaatg cccctcccc ttgcccttac 3240
tggaacttggc cctgcctgat gccaaagaaa gacccttccc cagaggccta cccccatat 3300

gtcctcagag aggctgagtg tccctccag gcagtcatgg gccctgaggc ccctcctgcc 3360
 tggccctgct cccagtgagg gaggtgactg tgtttccag agtgtgagcc gctctcctcc 3420
 ccctaaaaag ctgactcact gtgagtgacc ttgggcaagt tcccaaacct ccttgtgcct 3480
 cagtttcccc atctggaaaa aatggggcca cctcttgcca gcagtagcag ggctgcccac 3540
 gcccttttct ccccatgccc catccagcac ttgggcgact catgcctctg cctcagtggg 3600
 cctgtgggag cctactggag ccagcactt actccccctg agcagcgagc ctgcgtctgt 3660
 ctcagctgtc cagcgctgag ggccagggtc ttgtgtgtgt gggctggggg atgccctctt 3720
 ttctatatatt atttcataga aagtctcctg cgggagcgga aatgcagtcc ggcctagggc 3780
 tcccagccct tgactgtcct cctgtgaggg cctgaagctg ggccagggcc cgtcgagcg 3840
 gagccccctc tcagcagccc accgggtccc tccaggctgc tgcccgtgcg tggtctttct 3900
 cctccttttc aaagcaatag ccgccgggtc tgcaaagccc tgtcagacag actgggacct 3960
 tccaaggtca agccatgtgt ctgatgacat tcctggtgaa gcaaaggaga ggaggatggg 4020
 tcagccctca ctgggtgtca cacttgaga gaagtcctat tgtaaagaaa cggaaaaagt 4080
 cacaaaaaag tttgtataaa gacatatatt .tgtactacat ggggactctt cctgcatgtc 4140
 agcaataaaa cttcctgatc tggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaag 4199

<210> 29

<211> 3056

<212> DNA

<213> Homo sapiens

<400> 29

cttttctttg aatgacagaa ctacagcata atgcgtggct tcaacctgct cctcttctgg 60
 ggatgttgtg ttatgcacag ctgggaaggg cacataagac ccacacggaa accaaacaca 120
 aagggttaata actgtagaga cagtaccttg tgcccagctt atgccacctg caccaataca 180
 gtggacagtt actattgcgc ttgcaaaca ggcttcctgt ccagcaatgg gcaaaatcac 240
 ttcaaggatc caggagtgcg atgcaaagat atcaatgagt gcctcaccag cagcgtctgc 300
 cctgagcatt ctgactgtgt caactccatg ggaagctaca gttgcagctg tcaagttgga 360

ttcatctcta gaaactccac ctgtgaagac gtggatgaat gtgcagatcc aagagcttgc 420
ccagagcatg caacttgtaa taacactgtt ggaaactact cttgtttctg caaccagga 480
tttgaatcca gcagtggcca cttgagtttc caggggtctca aagcatcgtg tgaagatatt 540
gatgaatgca ctgaaatgtg ccccatcaat tcaacatgca ccaacactcc tgggagctac 600
ttttgcacct gccaccctgg ctttgcacca agcaatggac agttgaattt cacagaccaa 660
ggagtggaaat gtagagatat tgatgagtgc cgccaagatc catcaacctg tggtcctaata 720
tctatctgca ccaatgccct gggctcctac agctgtggct gcattgcagg ctttcatccc 780
aatccagaag gctcccagaa agatggcaac ttcagctgcc aaagggttct cttcaaatgt 840
aaggaagatg tgatacccga taataagcag atccagcaat gccaaagagg aaccgcagtg 900
aaacctgcat atgtctcctt ttgtgcacaa ataaataaca tcttcagcgt tctggacaaa 960
gtgtgtgaaa ataaaacgac cgtagtttct ctgaagaata caactgagag ctttgtccct 1020
gtgcttaaac aaatatccac gtggactaaa ttcaccaagg aagagacgtc ctccctggcc 1080
acagtcttcc tggagagtgt ggaaagcatg aactggcat ctttttgga accctcagca 1140
aatgtcactc cggctgttcg gacggaatac ttagacattg agagcaaagt tatcaacaaa 1200
gaatgcagtg aagagaatgt gacgttggac ttggtagcca agggggataa gatgaagatc 1260
gggtgttcca caattgagga atctgaatcc acagagacca ctggtgtggc ttttgtctcc 1320
tttgtgggca tggaatcggg tttaaagtag cgcttcttcc aagaccacca ggctcccttg 1380
accacctctg agatcaagct gaagatgaat tctcgagtgc ttgggggcat aatgactgga 1440
gagaagaaag acggcttctc agatccaatc atctacactc tggagaacgt tcagccaaag 1500
cagaagtttg agaggcccat ctgtgtttcc tggagcactg atgtgaaggg tggaagatgg 1560
acatcctttg gctgtgtgat cctggaagct tctgagacat ataccatctg cagctgtaat 1620
cagatggcaa atcttgccgt tatcatggcg tctggggagc tcacgatgga cttttccttg 1680
tacatcatta gccatgtagg cattatcacc tccttgggtg gcctcgtctt ggccatcgcc 1740
acctttctgc tgtgtcgtc catccgaaat cacaacacct acctccacct gcacctctgc 1800
gtgtgtctcc tcttggcgaa gactctcttc ctcgccggtg tacacaagac tgacaacaag 1860
atgggctgcg ccatcatcgc gggcttcttg cactacctt tccttgccctg cttcttcttg 1920
atgctgggtg aggctgtgat actgttcttg atggtcagaa acctgaagg ggtgaattac 1980
ttcagctctc gcaacatcaa gatgctgcac atctgtgcct ttggttatgg gctgccgatg 2040
ctggtgggtg tgatctctgc cagtgtgcag ccacagggt atggaatgca taatcgctgc 2100

tggctgaata cagagacagg gttcatctgg agtttcttgg ggccagtttg cacagttata 2160
 gtggtaagca aatactacaa cagcctggcg aagtgtgttc tgaaggagga gcaaggagac 2220
 ctgcgagatc tggaatttcc agggacgtgt gcagctgaga ggatcaactc ctttctcctg 2280
 acctggacct tgtggatcct gaggcagagg ctttccagtg ttaatgccga agtctcaacg 2340
 ctaaaagaca ccaggttact gaccttcaag gcctttgccc agctcttcat cctgggctgc 2400
 tcctgggtgc tgggcatttt tcagattgga cctgtggcag gtgtcatggc ttacctgttc 2460
 accatcatca acagcctgca gggggccttc atcttctca tccactgtct gctcaacggc 2520
 caggtacgag aagaatacaa gaggtggatc actgggaaga cgaagcccag ctcccagtc 2580
 cagacctcaa ggatcttgct gtcttccatg ccatccgctt ccaagacggg ttaaagtcct 2640
 ttcttgcttt caaatatgct atggagccac agttgaggac agtagtttcc tgtaggagcc 2700
 taccctgaaa tctcttctca gcttaacatg gaaatgagga tcccaccagc cccagaaccc 2760
 tctggggaag aatgttgggg gccgtcttcc tgtggttgta tgcactgatg agaaatcagg 2820
 cgtttctgct ccaaacgacc attttatctt cgtgctctgc aacttcttca attccagagt 2880
 ttctgagaac agaccctaat tcaatggcat gaccaagaac acctggctac cattttgttt 2940
 tctcctgccc ttgttggtgc atggttctaa gcgtgcccct ccagcgccta tcatacgcct 3000
 gacacagaga acctctcaat aaatgatttg tcgcctgtct gactgattta tcctag 3056

<210> 30

<211> 3155

<212> DNA

<213> Homo sapiens

<400> 30

tgatgaggaa gggttttaag gtaaggatgt ggtgacctag agtcttgttt ggtcctgagc 60
 accctctgct ggagcagagt aaaagtgact gacagttaa tcttggtttc tctacccag 120
 ccaactccag taaggctctc ttactcttg tgtgtttac acattgaagg catgttctat 180
 actgagagag aatagagaaa ccaaacagac tgacaggaca tggtaagaca gcatagaaaa 240
 acagaagtta gtggtttcgg ctgagggtac agtctttctg gtctgatgat tatatatagg 300

aatctattag ccaatgcctt aactattttc ctctcttccc aaatgaatga tgagactcac 360
ccttcatttt gatcagtttt aaggattttac ttcattctct gttcagtcct ctgtgtataa 420
agctgaagca gtagaggaag agagaaactc caagtgggtg catgactgct ctccactatt 480
actgtgccgt tgaagctgcc agtcgtcatg gaaataggct taaagttttt ctgcctcatc 540
tcaccatata tcaacagatc cagaccagcc ctggggcagg aactgatgat actgctagtg 600
cagacctagc tcaaactctc tgttctttca cagacatgtc tgctctgatt tcaactcagga 660
aacgtgctca gggggagaag cccttggctg gtgctaaaat agtgggctgt acacacatca 720
cagcccagac agcgggtgagt ttgttggag aaaagaggga cttctgataa ggggaagaaa 780
gaaatcaatt tttttttttt cacttttctt ttcctaactc gccaacctcc aaatgctaac 840
tattaatagg atctccctgt tacacaataa tggggttttc tgggtttttt ttttaatatg 900
taaaccact ttcagcaaac ttgacctagg aaagtctaga ttgcaaaca tatgattgtt 960
ataaaatgtt ataaagactt ctagaaacct atgttttatg caaagttaa taccagtgca 1020
ccacctcgct ggagtacacc cttttgtaac tttccagtgt agcacctgtt ggtgcagagt 1080
tgataggat tgctatcttg atgaggcttg ctctccattc ctgggtggcc cttctcttcc 1140
tcacctgtc ttctacact tctacacagg tgttgattga gacactctgt gccctggggg 1200
ctcagtgccg ctggtctgct tgtaacatct actcaactca gaatgaagta gctgcagcac 1260
tggctgaggc tggagttgca gtgttcgctt ggaagggcga gtcagaagat gacttctggt 1320
ggtgtattga ccgctgtgtg aacatggatg ggtggcaggc caacatgatc ctggatgatg 1380
ggggagactt aaccactgg gtttataaga agtatccaaa cgtgtttaag aagatccgag 1440
gcattgtgga agagagcgtg actggtgttc acaggctgta tcagctctcc aaagctggga 1500
agctctgtgt tccggccatg aacgtcaatg attctgttac caaacagaag tttgataact 1560
tgtactgtg ccgagaatcc attttggatg gcctgaagag gaccacagat gtgatgtttg 1620
gtgggaaaca agtgggtgtg tgtggctatg gtgaggtagg caagggtgc tgtgtgctc 1680
tcaaagctct tggagcaatt gtctacatta ccgaaatcga cccatctgt gctctgcagg 1740
cctgcatgga tgggttcagg gtggtaaagc taaatgaagt catccggcaa gtcgatgtcg 1800
taataacttg cacaggaaat aagaatgtag tgacacggga gcacttggat cgcataaaaa 1860
acagtgtgat cgtatgcaat atgggccact ccaacacaga aatcgatgtg accagcctcc 1920
gcactccgga gctgacgtgg gagcgagtac gttctcaggt ggaccatgtc atctggccag 1980
atggcaaacg agttgtcctc ctggcagaga ctcagatgga tttgcagaaa ataggaaaaa 2040

gcctagagtc ctgatctgga actgtaattc ttcctcttcc aactgagtct agaggggtgt 2100
 aaagtcctct cccgcttttg cttgacttgt atgccattgg aatctgagac agagctcatc 2160
 ccagggctct ctcttacctt tccagggtcg tctactcaat ttgagctgct ccacagttcc 2220
 cacctttgtt ctgtccatca cagccacaac acaggctttg gcactgatag aactctataa 2280
 tgcacccgag gggcgataca agcaggatgt gtacttgctt cctaagaaaa tggatgaata 2340
 cgttgccagc ttgcatctgc catcatttga tgcccacctt acagagctga cagatgacca 2400
 agcaaaatat ctgggactca acaaaaatgg gccattcaaa cctaattatt acagatacta 2460
 atggaccata ctaccaagga ccagtccacc tgaaccacac actctaaaga aatatttttt 2520
 aagataactt ttattttctt cttactcctt tcctcttgat ttttttcta taatttcatt 2580
 cttgtttttt catctcatta tccaagttct gcagaccaca caggaacttg cttcatggct 2640
 ctttagatga aatagaagtt cagggttcct cactctagtc actaaagaag gattttactc 2700
 tcccagccca gaaaggtgat tctttcttta ccatttctgg ggactttagt ctttaattagg 2760
 taccttatta acaggaaatg ctaaggtacc ttctctgtgg aacaatctgc aatgtctaaa 2820
 tcgccttaaa agagcccatt tctcagctgc tgaaatcagt gctctttcac ctcttcagag 2880
 aagcagggat ggtacctacc cggcaggtag gttagatgtg ggtggtgcat gttaatttcc 2940
 cttagaagtt ccaagccctg tttcctgcgt aaaggtggta tgtccagttc agagatgtgt 3000
 ataatgagca tggcttggtta agatcaggag gccacattgg atttatagta tagcccttcc 3060
 tccactccca ccagacttgc tcatttttcg agtttttaac tagactacac tctattgagt 3120
 ttaattttgt cctctaggat ttatttctgt tgtcc 3155

<210> 31

<211> 2626

<212> DNA

<213> Homo sapiens

<400> 31

cagtttatag cctggagtct ctgtttccct ttcaggtecc ggctctcttg gctttcccat 60
 aagcagctgc cttgggactc tcctggagac ctgatgccca cagccaagct gaccacagga 120

gccggtgctg gggactgagg gaaacttaga gttcagagag ggggtgtgat ttgcctgagg 180
tcacacagca agttagagac ccagctccac gactcattgt cttggctttg gccctcgtca 240
tcctgcccac ccagcggggc ttcccaaccc accacacagc catggacggg aaggtggcag 300
tgcaagagtg tgggcctcct gcagtctcct ggggtccccga ggaggagag aagttggacc 360
aggaagacga ggaccaggtg aaggatcggg gccaatggac caacaaggtg gagtttgtgc 420
tgtcagtggc cggggagatc attggtctgg catctgtggg catcgagtca tatttgaatg 480
tctactacat catcatcctt gcctgggctc tcttctacct gttcagctcc ttcacctctg 540
agctgccctg gacgacctgc aacaactttt ggaacacaga gcattgcacg gactttctga 600
accactcagg agccggcaca gtgaccccat ttgagaattt tacctcacct gtcattggaat 660
tctgggagag acgagttctg ggcatacct cgggcatcca tgacctgggc tccctgcgct 720
gggagctggc cctgtgcctc ctgctcgctt ggggtcatctg ctatttctgc atctggaagg 780
gggtcaagtc cacaggcaag gtggtttatt tcacagccac gtttccgtac ctgatgcttg 840
tcattttgct gatcagaggt gtcacccttc ccggagccta ccagggcata atctactact 900
tgaagccaga tttgttccgc ctcaaggacc ctcaggtgtg gatggatgcg ggcaccacaga 960
tcttcttctc ctttgccatc tgccaggggt gcctgacagc cctgggcagc tacaacaagt 1020
atcacaacaa ctgctacaag gactgcatcg ccctctgctt cctgaacagt gccaccagct 1080
ttgtggctgg gtttgttgc ttctccatcc tgggcttcat gtcccaagag caaggggtgc 1140
ccatttctga agtggccgag tcaggtcctg ggctggcctt catcgcttc cccaaggctg 1200
tgactatgat gcccttatcc cagctgtggg cctgcctgtt ctttatcatg ctcatattcc 1260
tagggctgga cagccagttt gtctgtgtgg agtgcctggg gacagcctcc atagacatgt 1320
tccccaggca gtcctggaag agcgggcggc gcgagctcct catcctcacc atcgccgtca 1380
tgtgctacct gatagggtt ttcttggtca ccgagggcgg gatgtacatc ttccagctgt 1440
ttgactacta tgcttccagt ggcatatgcc tgctgttcct gtcattgttt gaagtggctt 1500
gcataagctg ggtgtatggg gcggaccgtt tctatgacaa cattgaggac atgattggct 1560
ggttcctggc tctgtcctcc atgggtctgtg tcccactctt cgtcgtcatc accctcctga 1620
agactcgggg tcctttcagg aagcgtctgc gtcagctcat caccctgac tccagtctgc 1680
cacagcccaa gcaacatccc tgcttggtatg gcagtgtctg ccggaacttt gggccctccc 1740
caacaaggga aggactgata gccggggaga aggagacca tttgtagggg gtggccagag 1800
gccaggcggc tcctaagccg ggaacctagg tcagggccac cctccattct cagcggacag 1860

cctctgcctc tgtctcctgc cacaatcctg ctgggaacct ctggagagcc acaggcaccc 1920
 ccagctggag gccagactcc tctcttgtgc tagctggagc agctccttcc cctttgtga 1980
 taacaccacc actgggacgt gccatgttgg gacgccactc cctgtgggaa ggcaccatcg 2040
 tttttataaa ggggggtctt tttggaggcc gccatctgat tgcaaacact cgagttatga 2100
 ggattccact gtggggatgc ctcttgtag agcgtactgc atttgtacac ggggagagga 2160
 gctataattg gaacgcacac tgccgtccaa tgtggagagc ctgatgggac aataccctgt 2220
 tggaagtga aactgaacac actgtgttgg atcggagggt ccgttagggg atccttcctt 2280
 aggcctaacg acagaggcaa gcctttgcat gccgtcagtc tggagtttcc tccgagtctc 2340
 tcatggcatc tccagctcct gccctagtcc cgcactgttc ttgcagtgtt tcatcaactc 2400
 ctggagcatt ggaatggaag gggcttggga gatgattcct agacttcaca aacactcggc 2460
 atgcctccct gcaactgtccg ttccctctgcc caaggccgat attgctaact gatcacagat 2520
 tctttccac ctcacaatcc ttccgaatgt gctccaggca gcaccatttg ccatectgct 2580
 tctaacgcaa acccctgcct tagttaacaa gtgtgtatgg attgcc 2626

<210> 32

<211> 2673

<212> DNA

<213> Homo sapiens

<400> 32

ttacatactg cttttaggta gaaaagggga gggcagacag cacttcctgt atctgccatt 60
 tctccattgc cttcagctca aaataaccaa tttgccacag tggcatagtt tggagcaatg 120
 ctctgctcct cttcaaact tttccttate tcattcacca gtcagccacc aagtcctggt 180
 aacctcttct cttgtccaa actgtctact tctctccatc cccactgcca tgatctagtt 240
 ttgacctcca tcattcttct cttagtcttg tgactgggtct ctctgtctcc agtcataatc 300
 ttttttcagt ttattttcca tactgtagac agaactttct ttctaaaaca tgacagtatc 360
 acgttcctgc tcaaagctgc ccatgcactt ttcagttttc ttaaaacaaa acttaaacctc 420
 cttgacatgt aataaggcct gtgtggctgc acttttgtct tggctccaag gcttctcaag 480

tgtgggagcc tgtggaatcc tcagcttcct cacctgtgtg tgggaacagc agtacctttc 540
tcacagagtt gcccatggaa taagtggggc agtgcattga aagtgccttg catggtgtct 600
gatagtcaag gctgggtggc tgacggtttt caccctggat cttcagaaat ggtcatggcc 660
catccctcag aagactgctg aaaggactcc aggaagatg atgacatata gctcttgctc 720
ctctccttag gaaacccac tttatgtcca gtaagaactc ttcattggtga ggattaaagc 780
tcatttttct gtttcctata gaaacagcaa atgggggtgt ggtgggtgct aagggtgggtt 840
tctccctgaa aacctgtcca gctgtcgtag ggagggctgc tgagagcgtg gggccctgat 900
gcactatgag gatcagagta cctcctacat gccctggggc ccactgagag cttcatcctc 960
ccctgctcac tcacgggctc tgctccagct cagctgccat ccggcttcct gctcacagcc 1020
ttgcatctgt gtccttctct gcaggccaaa ctgcaccttt tctttccttc ctatcctctt 1080
cagttaatga aactggacc cagctgcttc tcattctatc cccctaaacc aatcctcctg 1140
gaaaatctcc tgtctccagg gtagtctgtg caccatagtg caatgatttt tcttcttggt 1200
ctgctcttgt gactcctgga acccttcgca gaccacagct gcaaacaaaa tcctcaactt 1260
ataatcacag agtgctacat aacagtcatt ggaaatgtct agtttaggga tgggtgtgtgt 1320
cctggatatct ttgcgtttat attcggaagg acttgggagc tctaaatata gggttcctgc 1380
aatggtaaac accatcaata gtactcatac ctatattacc ttcaggactg tagaaccttg 1440
gaaataatta cacttatgat ttgtctcagc tgtttaagtc aatgcctttt atttttagtt 1500
tttctgaaga caaagctctt ataagaatca cagatgaaag atcaggcaca aatcacattt 1560
tcccccttaa taacaaaata caaatccaat aattttagaa aatcagtttt tagtgacce 1620
gatgcctgga gaaaagctgc caggatTTTT ctggtctatc gcagaatttt ctacatcaat 1680
gagaaggatg ctgcatactt tggctgtatt atttctacc gtgagaaaag agacttagta 1740
tatggaacat gcttttttca gaaaattggc agtaactgac tttgaaggaa agttggttaa 1800
gttggacttg cagctggaac ttgggaagca ctgtccctc cttacccccg aggaaggaga 1860
cacagaggca cacttcagat aagtctcttg ttcagtgggt cactcatgtc ttcaacagcc 1920
agatctcatt gcgccgtccg tagggcttca tgggagggtc ataaccctg cagaagtaga 1980
tgtcccccg gtaggtggct gtgccctcca gggcagcacg cagacgggtg gcttgtgcta 2040
cgtagtctgc ttccttgga taaccaccaa actgcctggg ggttcaagag caagggtg 2100
ttacgaaagg acatgacctc tgagagaggg gggatcttta gaaaaaagtc actcaccatg 2160
tttattatga caagcatgct agagtggtaa gactggggac tgagtcacca cttcctagat 2220

gggttgattt tacctctctg agcctcagtt tcctcatctg aaaaaggaga tatattttct 2280
 ttttcctcta gagttttgag tgttaaaagc agtaatggac gtgaaaaact ctaatgcggg 2340
 gcctgaaaca gtagctgagt ataaacgtga gcctcccttt ctttcctgaa atacccaagg 2400
 ggacagcagg tttgctttcc caaaggccaa gccttcaccc ttttctctgg ctccctctac 2460
 tgcccggtg tacctctttt ctgctgctct gtcaccacct ctggaaggag gaagtctaca 2520
 ggggacttcg ggcagagctg cagtgagtat ggtggctaata gcccttcca cagcagagca 2580
 gctctagtga ctgagttcat cacagcagtg acttgtagca gcttcattta aaaacagagc 2640
 taggctagcc tgggcaacac aagaccctgt gtc 2673

<210> 33

<211> 3513

<212> DNA

<213> Homo sapiens

<400> 33

attttctccg ccggctgcgg gaaccggttc caggagcgcg ggagctcgtc tgccaggaga 60
 cccagccggg agggaccgat gctgcgtgat gcaaccctt cgagggactg agtcaagaaa 120
 ccaaggctgg gggctctctg gagtcaggcg gagctgggga ttaggactcg gggcttgtgt 180
 ctcccaggcg tgttttttt ttgttgtttt gtttttttgc cgcaacatct tccttatact 240
 ggggtggggtg ggggcggggg tggctctgtat ttccagacct cctccatggt ggtctggctg 300
 tgtggagggc agggcacagt tccttgactt ggctcattcc cccccaacgc ccccccttc 360
 tttttttttt gagatggagt ctcgctctgt tgcccaggct ggagtgcagt agcgtgatct 420
 cggctcactg caatctccgt ctcccgggtt caagcgaatc tcctgcctca gcctccagag 480
 tagctgggac tacaggcacc caccaccacg cccggctaata ttttgtattt ttagtagaga 540
 cagggtttca ccatattggc caggctggtc tggaactcct gaccttgtga tccgcctgcc 600
 tcggcctcct aaagtgtctg gattacaggc gtgagccacc gcgcctggca ggctcattcc 660
 ctttttatgc cttcgttttt gatagatgac ctggggatgg gagaaatgct gcagaggtgg 720
 tttatttaca gtgaattgtt aatttggttg tggggctagt attaggggca cattgggaaa 780

acacaggctg ggcagaggtt ggtaggttca caacagttat gccactggg tgctgatgga 840
tttgggggct gtcgtcggcc tggcttgggg ccctagtggg ggggctacaa ggcttgctcc 900
tagatctact ccaagtgaac atttaagatc agtgacttgg atggagactt aaaaaggaag 960
tcagctgact ccgtgtagtc atccagagct gtgaggaatg acaagtgcac agtataaaat 1020
caagattcta attgaaggcc tccaccatgg acagaggcca ggccctgccc ctcccaggca 1080
gcctggctcc ttctgctggg ccctgaaggc agacgggata atgtggttgg ccaaggcctg 1140
ttggtccatc cagagtgaga tgccctgtat ccaagcccaa tatgggacac cagcaccgag 1200
tccgggaccc cgtgaccacc tggcaagcga cccctgacc cctgagttca tcaagccac 1260
catggacctg gccagccccg aggcagcccc cgtgcccc actgccctgc ccagcttcag 1320
caccttcacg gacggctaca caggagagtt tgacaccttc ctctaccage tgccaggaac 1380
agtccagcca tgctcctcag cctcctcctc ggctcctcc acatcctcgt cctcagccac 1440
ctccccctgcc tctgcctcct tcaagttcga ggacttccag gtgtacggct gctaccccg 1500
ccccctgagc ggcccagtgg atgaggccct gtctccaggt ggctctgact actatggcag 1560
cccctgctcg gccccgtcgc cctccacgcc cagcttccag ccgccccage tctctccctg 1620
ggatggctcc ttcggccact tctcgcccag ccagacttac gaaggcctgc gggcatggac 1680
agagcagctg cccaaagcct ctgggcccc acagcctcca gccttctttt ccttcagtcc 1740
tcccaccggc ccagcccca gcctggccca gagccccctg aagttgttcc cctcacaggc 1800
caccaccag ctgggggagg gagagagcta ttccatgcct acggccttcc caggtttggc 1860
accacttct ccacaccttg agggctcggg gatactggat acaccctga cctcaaccaa 1920
ggccccggagc ggggccccag gtggaagtga aggccgctgt gctgtgtgtg gggacaacgc 1980
ttcatgccag cattatggtg tccgcacatg tgagggtgc aagggttct tcaagcgcac 2040
agtgcagaaa aacgccaagt acatctgcct ggctaacaag gactgccctg tggacaagag 2100
gcggcgaaac cgctgccagt tctgccgctt ccagaagtgc ctggcggtgg gcatggtgaa 2160
ggaagttgtc cgaacagaca gcctgaaggg gcggcggggc cggctacctt caaaacccaa 2220
gcagccccc gatgcctccc ctgccaatct cctcacttcc ctggtccgtg cacacctgga 2280
ctcagggccc agcactgcca aactggacta ctccaagttc caggagctgg tgctgcccc 2340
ctttgggaag gaagatgctg gggatgtaca gcagttctac gacctgctct ccggttctct 2400
ggaggtcatc cgcaagtggg cggagaagat ccctggcttt gctgagctgt caccggtga 2460
ccaggacctg ttgctggagt cggccttctt ggagctcttc atcctccgcc tggcgtaacag 2520

gtctaagcca ggcgagggca agctcatctt ctgctcaggc ctggtgctac accggctgca 2580
 gtgtgcccgt ggcttcgggg accggattga cagtatcctg gccttctcaa ggtccctgca 2640
 cagcttgctt gtcgatgtcc ctgccttcgc ctgcctctct gcccttgctc tcatcaccga 2700
 ccggcatggg ctgcaggagc cgcggcgggt ggaggagctg cagaaccgca tcgccagctg 2760
 cctgaaggag cacgtggcag ctgtggcggg cgagccccag ccagccagct gcctgtcacg 2820
 tctgttgggc aaactgcccg agctgcggac cctgtgcacc cagggcctgc agcgcattct 2880
 ctacctcaag ctggaggact tgggtgcccc tccacccatc attgacaaga tcttcatgga 2940
 cacgctgccc ttctgacccc tgcctgggaa cacgtgtgca catgcgcaact ctcatatgcc 3000
 accccatgtg ccttttagtcc acggaccccc agagcacccc caagcctggg cttgagctgc 3060
 agaatgactc caccttctca cctgtctccag gaggtttgca gggagctcaa gcccttgggg 3120
 aggggggatgc cttcatgggg gtgacccac gatttgtctt atcccccca gcctggcccc 3180
 ggcctttatg ttttttgtaa gataaacctt ttttaacaca tagcgccgtg ctgtaaataa 3240
 gcccagtgtc gctgtaaata caggaagaaa gagcttgagg tgggagcggg gctgggagga 3300
 agggatgggc cccgccttcc tgggcagcct ttccagcctc ctgctggctc tctcttctta 3360
 cctccttcc acatgtacat aaactgtcac tctaggaaga agacaaatga cagattctga 3420
 catttatatt tgtgtatatt cctggattta tagtatgtga cttttctgat taatatattt 3480
 aatatattga ataaaaata gacatgtagt tgg 3513

<210> 34

<211> 3435

<212> DNA

<213> Homo sapiens

<400> 34

ttccgggttt ggtgcggcct ttgtctctcg ctgcagtcag agctccaggt ctggttctta 60
 cttctcctaa aggcccaggc tgtgtggccc cgtgtcctgc aggtattggg agatccacag 120
 ctaagacacc gggacctcct ggaagccaaa aatgggacca ttgcaattta gagatgtggc 180
 catagaattc tctctggagg agtggcattg cctggacact gcacagcgga atttatatag 240

gaatgtgatg ttagagaact acagtaacct ggtcttcctt ggtattactg tttctaagcc 300
agacctgatc acctgtctgg agcaaggaag aaaacctttg accatgaaga gaaatgagat 360
gatagccaaa ccctcagtaa tgtgtttctca ttttgcccaa gacctttggc cagagcagag 420
catgaaagat tctttccaaa aagtgggtact gagaagatat gaaaaatgtg aacatgacaa 480
tttacagtta aaaaaaggat gtataagtgt ggatgagtgt aagggtgcaca aagaaggtta 540
taatgaactt aaccaatgtt tgacaactac cccaagaaaa atatgtcaat gtgataaata 600
tgtgaaagtc cttcatcaat ttccaaattc aaacggacaa aagagaggac atactgggaa 660
aaaacctttc aaatatatag aatgtggcaa agctttttaag cagtttctca ctcttactac 720
acataagaaa attcatactg gagggaaacc ctacaaatgt gaagaatgtg gcaaagcctt 780
taaccactct ttagcctta ctagacataa gaaaattcat actggagaga aaccctacaa 840
atgtgaagaa tgtggcaaag cttttaagca ctctccact cttactacac ataagagaaa 900
tcatactgga gagaaaccct acaagtgtga taaatgtggc aaagccttta tgtcatcctc 960
aacccttagt aaacatgaga taattcatac ggaaaagaaa ccctacaaat gtgaagaatg 1020
tggcaaagcc ttcaaccgtt cctcaaccct tactacacat aagataattc atactggaga 1080
aaaaccctac aaatgtgaag aatgtgacaa agcctttaag tactcctata cccttactac 1140
acataaaaaga attcatactg aagacaaacc ctacaaatgt gaagaatgtg gcaaagcctt 1200
taagtactcc tctaccctta ctacacataa gagaattcat actggagaga aaccctacaa 1260
atgtgaagaa tgtggcaaag ctttcaagcg gtcttcagac cttactacgc ataagataat 1320
tcatactgga gagaaaccct acaaatgtga agaatgtggc aaagccttta agtactcctc 1380
taaccttact acacataaga aaattcatac tggagagaga ccctacaaat gtgaagaatg 1440
tggcaaagcc ttcaaccagt cctcaatcct tactacacat aggagaattc atactggaga 1500
gaaattctac aaatgtgaag aatgtggcaa agctttttaag tgctcctcta accttactac 1560
acataagaaa attcatactg gagagagacc ctacaaatgt gaagaatgtg gcaaagcctt 1620
caaccagtcc tcaatcctta ctacacatga gagaatcata ctggagagaa attctacaaa 1680
tgtgaagaat gtggcaaagc cttcaagcgg tcctcacacc ttactacaca taagataatt 1740
catactggag agaaattcta caaatgtgaa gaatgtgaca cagcttttaa gcatttctct 1800
aaccttacta cacataagaa aattcatact ggaaaaaac cctacagatg tgaagaatgt 1860
ggcaaagcct ttaagcactc ctctaaactt actacacata agagatttca tactggagag 1920
aaaccctaca gatgtgaata atgtggtaaa gcttttcacc tctcctcaca ccttactgca 1980

cataagataa ttcatactgg agagaaacca tgtgaatgtg atgaatgggg aaaatctttt 2040
aaccagccct caatttctaa ctaagaattg atatgaaata taaactctac aaatataaag 2100
aatgtgacag acggacactc catctcaaaa agaaagaaaa aagccaggga ttacaggcgt 2160
gagccaccgc gtccggcctc tttttttctt ttcttttttt tgagacaaag tctcactgtg 2220
tcaccagac tggaatgcag tgacacaatc tcggctcact gaaacctctg ccttcaggt 2280
tcaagctatt ctcatgcctc agcctctcaa gtagctggga ctacagatgt gggccaccat 2340
gtctggctaa tttttttttt tttttttttt ttgtagagac agggtttcgc catgttgacg 2400
agactggctc cgaactcctg gcctcaagca atccaccac ctcaggctct caaagtgttg 2460
ggattatagg tgtgagccaa tatgtccagc ctaattttat tttatttttt accatcttaa 2520
aactattgtg catttaatga agcattatta tgccactaac tttaacctat ccggccttac 2580
tcaaattggtg taagcaaaag atggaagaag tatactatit ggtatgtagt ggaatgtcat 2640
cactagtaat cactttgcca gtaaactgca aatgaagaat attgttcccg ttagttaaat 2700
tttcaatctt ttttcttaat ttattattac tatctgtgag tatatagtat gtgtatatat 2760
gtcttataga gaatattttg attcaggcat acaatatgta gtaattacat tagggtaaata 2820
aaggatctg tcacctctag catttatect ttctaatacca aacagtgtaa ttatacactt 2880
ttagttactt ttaaagtgtc aattaaattg taattgacta cagggtgatt tttatggtca 2940
taataaaaaat tatacagaaa tataaataga atatgggctg gacgcagtgg ctcatgcctg 3000
taatcccagc actttggtag gcgacgtcgg gcagatcatg aggtcaggag ttcaagacca 3060
gcctgtcctg ggccagggat ctctctacaa gagcccctgc ccctctgttg gaggcacagt 3120
tttagaataa ggaggaggag ggagaagaga aaatgtaaag gagggagatc tttcccaggc 3180
cgcaccattt ctgtcactca catggaccca agataaaaaga atggccaaac cctcacaacc 3240
cctgatgttt gaagagttcc aagtigaagg gaaacaaaga agtgtttgat ggtgccagag 3300
aggggctgct ctccagaaag ctaaaattta atttcttttt tcctctgagt tctgtacttc 3360
aaccagccta caagctggca cttgctaaca aatcagaaat atgacaatta atgattaaag 3420
actgtgattg ccacc 3435

<210> 35

<211> 4932

<212> DNA

<213> Homo sapiens

<400> 35

atgatcatgt	cccctgtcca	cctgctccag	ccactatccc	tctcccactt	acagcagaag	60
aaagggctgg	tgagaaaggt	ggattacagg	cccacttctg	ccactgacga	gccctatgaa	120
tgtggcctac	acccccttag	cttcaactggg	tctcagtttc	cctatctgta	tattgggagc	180
agttgtgaag	ctcagaagag	aaatgtctgt	gaaaagggtta	tgaacaggag	ggagagtgga	240
aaccaacctg	ctggatcgtg	tccacagacc	ctggaatggg	gccacatgct	tggtttgtca	300
aattgcagac	gccggccggg	tgcgatggct	catgcctgta	atcccagcac	tttgggaggc	360
cgaggcggac	agatcacttg	aggctcgggag	ttcgagacca	gcctgaccaa	catggagaaa	420
ccccgtctct	actgaaaata	caaaattagc	caggcatggg	ggcacatgcc	tataatccca	480
gctacttggg	aaggctgagg	caggagaatc	acttgaacct	gggagacgga	ggttgtggtg	540
agcctagatc	gtgccattgt	actccagcct	gggcaacaag	agtgaaactc	cgtctcaaaa	600
aaaaaaaaatt	tgcagacgcc	atcccatcca	ggcctttgct	ttcactgatg	aagaaactga	660
gatacagaga	gggcagggca	cctgttcgga	gtttatgaaa	tgcccccca	ccattatctt	720
tcttgatcat	ataagaatct	ggtgaggcaa	ggtagggcgt	gatctttatc	tctattttat	780
cgttttat	aagcgggaac	aggactgctc	agtggctggg	ggccttgccc	aagatctcca	840
agtactgggg	aaccccaggg	aggccctggg	gggtggcagt	gttcctat	cagccccact	900
ctgcttcccc	ctcccaggat	atccaggagg	tgcagggcta	cgtgctcatc	gtcacaacc	960
aagtgaggca	ggtcccactg	cagaggctgc	ggattgtgcg	aggcaccag	ctctttgagg	1020
acaactatgc	cctggccgtg	ctagacaatg	gagaccgct	gaacaatacc	accctgtca	1080
caggggcctc	cccaggaggc	ctgcgggagc	tgcagcttcg	aagcctcaca	gagatcttga	1140
aaggaggggt	cttgatccag	cggaaccccc	agctctgcta	ccaggacacg	at tttgtgga	1200
aggacatctt	ccacaagaac	aaccagctgg	ctctcacact	gatagacacc	aaccgctctc	1260
gggcctgcc	ccctgttct	cgatgtgta	agggtcccg	ctgctgggga	gagagtcttg	1320
aggattgtca	gagcctgacg	cgcactgtct	gtgccgggtg	ctgtgcccgc	tgcaaggggc	1380
cactgcccac	tgactgctgc	catgagcagt	gtgctgccgg	ctgcacgggc	cccaagcact	1440
ctgactgcct	ggcctgcctc	cacttcaacc	acagtggcat	ctgtgagctg	cactgcccag	1500

ccctggtcac ctacaacaca gacacgtttg agtccatgcc caatcccagag ggccggtata 1560
cattcggcgc cagctgtgtg actgcctgtc cctacaacta cctttctacg gacgtgggat 1620
cctgcaccct cgtctgcccc ctgcacaacc aagaggtgac agcagaggat ggaacacagc 1680
ggtgtgagaa gtgcagcaag ccctgtgccc gagtgtgcta tggctctgggc atggagcact 1740
tgcgagaggt gagggcagtt accagtgcc aatatccagga gtttgctggc tgcaagaaga 1800
tctttgggag cctggcattt ctgccggaga gctttgatgg ggaccagcc tccaacactg 1860
ccccgctcca gccagagcag ctccaagtgt ttgagactct ggaagagatc acaggttacc 1920
tatacatctc agcatggccg gacagcctgc ctgacctcag cgtcttccag aacctgcaag 1980
taatccgggg acgaattctg cacaatggcg cctactcgtc gacctgcaa gggctgggca 2040
tcagctggct ggggctgcgc tcaactgagg aactgggcag tggactggcc ctcatccacc 2100
ataacacca cctctgcttc gtgcacacgg tgccctggga ccagctcttt cggaacccgc 2160
accaagctct gctccacact gccaacccgc cagaggacga gtgtggtaag acaggagacc 2220
cagtgtgcgc actccccatc tgccagcaca cagcagtgcc cagggggccc tggcagcagc 2280
gttcttggac ttgtgcagac tgcccgtctc tgtgcaccct tcttgactca gcacagctct 2340
ggctggcttg gcctcttggc atggcttctc tagctgggtc ctacctgcct tggcatcctt 2400
ccctccccct ctgtttctga aatctcagaa ctcttcctct ccctacatcg gccccacctg 2460
tccccacccc tccagcccac agccatgccc acagccagtt ccctggttca cttggacctg 2520
gggcctcccc taaaagtccc ctgcgggtccc ttctctctca ctgcagtggg cgagggcctg 2580
gcctgccacc agctgtgcgc ccgagggcac tgctggggtc cagggcccac ccagtgtgtc 2640
aactgcagcc agttccttcg gggccaggag tgcgtggagg aatgccgagt actgcagggg 2700
tatgaggggc ggaggagagg gtggctggag gggtgcatgg ggctcctctc agacccctc 2760
accactgtcc cttctctcag gctccccagg gagtatgtga atgccaggca ctgtttgccg 2820
tgccaccctg agtgtcagcc ccagaatggc tcagtgcact gttttggacc ggaggctgac 2880
cagtgtgtgg cctgtgccc aataaaggac cctcccttct gcgtggcccc ctgccccagc 2940
ggtgtgaaac ctgacctctc ctacatgccc atctggaagt ttccagatga ggagggcgca 3000
tgccagcctt gccccatcaa ctgcaccac tcctgtgtgg acctggatga caagggtgc 3060
cccgccgagc agagagccaa ccctctgacg tccatcgtct ctgcggtggg tggcattctg 3120
ctggctgtgg tcttgggggt ggtctttggg atcctcatca agcgacggca gcagaagatc 3180
cggaagtaca cgatgcggag actgctgcag gaaacggagc tgggtggagcc gctgacacct 3240

agcggagcga tgcccaacca ggcgcagatg cggatcctga aagagacgga gctgaggaag 3300
gtgaaggtgc ttggatctgg cgcttttggc acagtctaca agggcatctg gatccctgat 3360
ggggagaatg tgaaaattcc agtggccatc aaagtgttga gggaaaacac atccccaaa 3420
gccaacaag aaatcttaga cgaagcatac gtgatggctg gtgtgggctc cccatatgtc 3480
tcccgcttc tgggcatctg cctgacatcc acggtgcagc tggtagacaca gcttatgccc 3540
tatggctgcc tcttagacca tgtccgggaa aaccgcggac gcctgggctc ccaggacctg 3600
ctgaactggt gtatgcagat tgccaagggg atgagctacc tggaggatgt gcggctcgta 3660
cacagggact tggccgctcg gaacgtgctg gtcaagagtc ccaacctatg caaaattaca 3720
gacttcgggc tggctcggct gctggacatt gacgagacag agtacctatg agatgggggc 3780
aaggtaggt gaaggaccaa ggagcagagg aggctgggtg gagtggtgtc tagcccatgg 3840
gagaactctg agtggccacc tccccacaac acacagttgg aggacttcct cttctgcct 3900
cccagggtgcc catcaagtgg atggcgctgg agtccattct ccgccggcgg ttcaccacc 3960
agagtgatgt gtggagttat ggtgtgtgat ggggggtgtt gggaggggtg ggtgaggagc 4020
catggctgga gggaggatga gagctgggat ggggagaatt acggggccac ctcagcatgt 4080
gaaggaggagg aaggggctgc ctgtgcccc ccttgcaggg tctgtgact tcccaggatt 4140
agggaaagac cgggtagggt ctgtctcctg gcatcacatc tccccctgct acctgccatg 4200
atgctagact cctgagcaga acctctggct cagtacacta aagctccctc tggccctccc 4260
actcctgacc ctgtctctgc cttagggtgtg actgtgtggg agctgatgac ttttggggcc 4320
aaaccttacg atgggatccc agcccgggag atccctgacc tgctggaaaa gggggagcgg 4380
ctgccccagc ccccatctg caccattgat gtctacatga tcatggtcaa atgtgcgtgg 4440
ctgagctgtg ctggctgcct ggaggagggt gggaggctcct gggtggagga gccacaagg 4500
ggcatgaaag gggaccagga tgtatgtaga cccaggagcc ctagtatgtt aggagcctca 4560
aaaccttctt gtatcccttt tacagtcaaa gtccaaagcc actcttgagg aacactcttg 4620
tacaaaatta agctgggcac agtggctcat gcctgtaatc ccagtacttt tggaggctga 4680
ggtgggagga tcccttgaag ccaggagttc aagaccagcc tgggcaacat agtgagatcc 4740
tatctctaca aaaaataaaa aaattatctg ggtgtggtgg tgtgtgccag tagtcccagc 4800
tactcaggag aggctgaggc aggaagatca cttgagccta gtttaaggtt gcagtaagct 4860
atgattgcac cactgaaatc cagcctgggt gacagagcga aacctatct caaaaaata 4920
aaaaagcaaa ac 4932

<210> 36

<211> 2404

<212> DNA

<213> Homo sapiens

<400> 36

```
cgttcccttcg cacgtgcctc tgcgggattc ctcgtgggaa agtctcagaa atgtccctaa    60
actgaagaaa agccggggaa tgcaaggaga gcaaggctcg aggttggtcc gactgtcctc    120
taacgacctg cctggcaggg accgactacc ctgtccttac ccgcagcccc atgtggagca    180
ggcatccacc agtgcctccc cgcgcttgct agcagtgaac attcctgaag acactctcct    240
ctcttatact gaatcctcgc tccactgca gagacactgc gccttccata ggttctccaa    300
ggccaaatct gagtccctct ttaggacgta tgacaaggac atcaccttc agtatcttaa    360
gagcttcaaa cgagtcagaa taaacttcag caacccttc tccgcagcag atgccaggct    420
ccagctgcat aagactgagt ttctgggaaa ggaaatgaag ttatatattg ctcagacctt    480
acacatagga agctcacacc tggctccgcc aaatccagac aagcagtttc tgatctcccc    540
tccgcctct cccgagtggt gatggaaaca agtggaagat gcgacccag tcataaacta    600
tgatctctta tatgcatca ccaagctggg gccaggggaa aagtatgaat tgcacgcagc    660
gactgacacc actcccagcg tgggtgggtcca tgtatgtgag agtgatcaag agaaggagga    720
agaagaggaa atggaaagaa tgaggagacc taagccaaaa attatccaga ccaggaggcc    780
ggagtacacg ccgatccacc tcagctgaac tggcacgcga cgaggacgca ttccaaatca    840
tactcacggg aggaatcttt tactgtggag gtggctggtc acgacttctt cggaggtggc    900
agccgagatc ggggtggcag aaatcccagt tcatgttgct cagaagagaa tcaaggccgt    960
gtcccccttg tctaagtctg cacaccagtt actgttcatg gcacccggga atgacttggg   1020
ccaatcactg agtttgtggt gatcgcaaaa ggacatttgg gactgtcttg agaaaacaga   1080
taatgatagt gttttgtact gtctcttttc tggtaggttc tgtctgtgcc aagggcaggt   1140
tgatcagtga gctcaggaga gagcttcctg tttctaagtg gcctgcaggg gccactctct   1200
actggtagga agaggtacca caggaagccg cctagtgcag agaggttgtg aaaacagcag   1260
```

caatgcaatg tggaaattgt agcgttttctt ttctttccctc atgtttctcat gtttgtgcat 1320
 gtatatattact gatttacaag actaacctttt gttcgtatat aaagttacac cgttggttgtt 1380
 ttacatctttt tgggaagcca ggaaagcggtt tggaaaacgt atcacctttc ccagattctc 1440
 ggatttctga ctctttgcaa cagcacttgc ttgcggaact cttcctggaa tgcattcact 1500
 cagcatcccc aaccgtgcaa cgtgtaactt gtgcttttgc aaaagaagtt gatctgaaat 1560
 tcctctgtag aatttagctt atacaattca gagaatagca gtttctactgc caacttttag 1620
 tgggtgagaa attttagttt aggtgtttgg gatcggacct cagtttctgt tgtttctttt 1680
 atgtggtggt ttctatacat gaatcatagc caaaaacttt tttggaaact gttggttag 1740
 atagttggtt cttttacccc acgaagacat caagatacac ttgtaaataa agctgatagc 1800
 atatattcat acctgttgta cacttgggtg aaaagtatgg cagtgggaga ctaagatgta 1860
 ttaacctacc tgtgaatcat atgttgtagg aaaagctggt cccatgtcta acaggacttg 1920
 aattcaaagc atgtcaagtg gatagtagat ctgtggcgat atgagaggga tgcagtcctt 1980
 ttccccattc attcctgatg gaattgttat actaggttaa catttgtaat ttttttctag 2040
 ttgtaatgtg tatgtctggt aaataggtat tatattttgg ccttacaata ccgtaacaat 2100
 gtttgtcatt ttgaaatact taatgccaag taacaatgca tgctttggaa atttggaaga 2160
 tggtttttatt ctttgagaag caaatatgtt tgcattaaat gctttgattg ttcatatcaa 2220
 gaaattgatt gaacgttctc aaacctgtt tacgggtactt ggtaagaggg agccggtttg 2280
 ggagagacca ttgcatcgct gtccaagtgt ttcttgtaa gtgcttttaa actggagagg 2340
 ctaacctcaa aatatttttt ttaactgcat tctataataa atgggcacag tatgctcctt 2400
 acag 2404

<210> 37

<211> 1670

<212> DNA

<213> Homo sapiens

<400> 37

tttggtttaa aatcgtgatc ccagattatt catttaaagc aagtatatgg ctttgttttt 60

tggtctgttt tttgtttttg agacagagtc tcgctctgtt accaaggctg gagtgcagtg 120
gcgcgatctc agttcattgc agcctccgcc tcctgagttc aagcaattat tgtgtcccag 180
tctcctgagt tgctgggatt acaggcacac aacactacat ccggctaaaa tttttgtatt 240
tttagtagag acagagtttc gccatgttgg ccaggctggg cttgaactcc tgaccccaag 300
tgatctgcct gccgcagtct ctcaaagtat tgggactaca ggcatgagcc accatgcctg 360
gctgaaagca agtatataac tttgaagagt atctggatac ttaacgttta atgagtgcac 420
tagacaagtt gtcattggaag caaagttaaa tcctaagtcc gagttggaag caccatgggt 480
ggagctggtc aatatactgc ctaacactgt ccagctatga taatctgcat ggatctctgc 540
taggagacct aaccagtcat ggggacatct ttgaataaac cagtcagggc agacgtacct 600
ttggacaatg ccagtgtctac acgaatgaat ttttttaatt ttttgtatta tagagacagg 660
gtctctgttg cccaggctgg agtgcagtga cacaattgta gctcactgca gcctgcaatt 720
cctgggttca agagaccctc ccacctcagt ctcccaagta gctgagacta caggcactca 780
ccacctcatc cagctaatat tttagtttgt agagatggag gtcttgctat gttgccagg 840
ccggtctcaa actcctggcc tcctgaagta ctgggattac aggtgtgagc cactgcactt 900
ggccctacac cgatgaattt tttttttttt ttttttatag ggacagagtc ttgctctgtc 960
acccaggggtg gagtgcagtg agctatgagg gtgcctctgt gctccatcct aggcaacaga 1020
gtgagacgct gtttaaaaag gaaaaaatcc ttccttagag ctagtatcct aaagctgcag 1080
agctagccca gacctattg gtttccttgt ccttgggggtg cttttcctga atctttgggg 1140
gtgaagggag tgttgctccc agtccagagg cctgattctg ttcggactgg gttctcaaga 1200
cacgaccagg ttctcaagac acgagtcctt ttgttcctcc ccattaaagg gggtttgtca 1260
gaagcaagaa cagccccctc cccagtcac agcctgaagg gaggccccga gagcttcctc 1320
cttccccca cctgtcctt acctctctg cctgtcttt tagaactgca gttcattgtt 1380
ttaagggatt gggggaggga gcctggggac acaaacctt tatacaatac aaagctttgc 1440
ttttttttt tttttcttc ttttccttt ctcggttctc ttctctctc tgaatggctg 1500
aagaccctc tgccgaggga ggttggggat tgtgggacaa ggtcccttgg tgctgatggc 1560
ctgaaggggc ctgagctgtg ggcagatgca gttttctgtg ggcttgggga acctctcacg 1620
ttgctgtgtc ctggtgagca gcccagcaaa taaacctgct tttctaaaag 1670

<210> 38

<211> 2682

<212> DNA

<213> Homo sapiens

<400> 38

```
cgctcgccct ggggccgctc agggagccgg gacgcgctgc ctgggctccg cgcagccgat 60
gacattctga gttcttctc tccataccgg gctccacgtc tcctgtcttc ttacctgat 120
ccacaggacc cctcaacttt cctgtaaggg aggcctctatt accccgtttt tcagagacga 180
gaggtagggc tcagagaggt ttagttactt tccaagatca cacagccagt cagggcagag 240
ccgggacaca accctgtgtc tctggcggtg gctctatgga ctgggccct ctgttcccca 300
ggccccacct cttccagaac ctctgtctct tcctgtgggc cctgctgaac tgtggtttgg 360
gggtcagtgc tcagggtccg ggcgagtgga ccccggtggg gtccctggacc cgctgctcca 420
gctcctgcgg gcgtggcgtc tccgtgcgca gccggcgctg cctccggctt cctggggaag 480
aaccgtgctg gggagactcc catgagtacc gcctctgcca gttgccagac tgccccccag 540
gggctgtgcc cttccgagac ctacagtgtg ccctgtacaa tggccgccct gtcctgggca 600
cccagaagac ctaccagtgg gtgcccttcc atggggcgcc caaccagtgc gacctcaact 660
gcctggctga ggggcacgcc ttctaccaca gcttcggccg cgtcctggac ggcaccgcct 720
gcagcccggg tgcccagggg gtctgcgtgg ctggccgctg ccttagcgcc ggctgtgatg 780
ggttgttggg ctcggtgcc ctcgaggacc gctgtggccg ctgcggaggc gccgacgact 840
cgtgcctttt cgtgcagcgc gtgtttcgtg acgccggtgc cttcgtggg tactggaacg 900
tgaccctgat ccccgagggc gccagacaca tccgcgtgga acacaggagc cgcaaccacc 960
tggcactgat ggggggcgat gggcgctacg tgcttaatgg gactgggtg gtcagcccac 1020
cagggacctc cgaggcggcc ggcacgcatg tggctctacac ccgagacaca gggccccagg 1080
agacattgca agcagccggg cccacctccc atgacctgct cctacaggtc ctctgcagg 1140
agcccaaccc tggcatcgag tttagattct ggctccctcg ggagcgctac agccccttcc 1200
aggctcgtgt gcaggccctg ggctggcccc tgaggcagcc tcagccccgg ggggtggagc 1260
ctcagcccc cgcagcccct gctgtcacc ctgcacagac cccaacgtg gccccagacc 1320
cctgcccacc ctgccctgac acccgcggcc gcgcccaccg actactccac tattgcggca 1380
```

gtgactttgt gttccaggcc cgagtgtggt gccaccacca ccaggcccag gagaccgct 1440
atgagggtgcg catccagctc gtctacaaga accgctcgcc actgcgggca cgcgagtacg 1500
tgtgggcgcc aggccactgc ccctgcccga tgctggcacc ccaccgggac tacctgatgg 1560
ctgtccagcg tcttgtcagc cccgacggca cacaggacca gctgctgctg cccacgccc 1620
gctacgcccg gccctggagc cctgcggagg acagccgcat acgcctgact gcccggcgct 1680
gtcctggctg agcccctgca ggagccccgg ccacacacag caagaaagat acatctgacc 1740
agcctcaacg tcaacgtatt tccccctca ccctggcttc caggcagctc tgaaatacgt 1800
cccacctgtg cagctatgtg actccctccc acacacgctt aagacacctc tgcattgcagt 1860
caaagccact gtcacaagcc ggcaggcact ggtgaggagg cactaaggag actctgactt 1920
ttatttcgcc tctctccttg gctgccagga agctcatagc tatttatact cagaaagttt 1980
aacgctgctt tctttctctt tgcgcgcgtc acacttgctt ggagacactg tcatgaacga 2040
gcatgacacc ctgctgccct gggtagccag aagatcatct gtttacttcc cagacactgt 2100
gctgtctctg ctctctgcta ctacacaca ccctcatgtg tgaagggcag agacactgtc 2160
acaaacaggc atgcccctta gaagacatgc ctaaccaggc actgtaacgt accaacgtac 2220
caatttcccc ttttccctg gctaccagga aactcggaga caatctttc agcctcagca 2280
tttctggctg gatttcacc catcaacacg tgcttgctcc tcctttttt ttttttctga 2340
ggtagaccct tgctctgtca cctaggctga agtgcggtgg tgcaatcatg gctcactgca 2400
gcctcaaatt cctgggctca agcgatcctc ccacctcagc ctccacagca gctggaactc 2460
cagggtgtgaa tcaccatgcc cggctaactt ttaaattttt tgtagagatg gggctcttgc 2520
acgttgccca cactgctctc caactactgg cctcaaacga tcctcctgcc tcggcctccc 2580
aaagcactgg aatttcaggc atgcactatc gtactggctg ttcaccccc ttttattcct 2640
ctttaatttt tacaagcact ttttaaaaaa aaattttact tt 2682

<210> 39

<211> 4315

<212> DNA

<213> Homo sapiens

<400> 39

tttctgtatt ccgggttctt gccttggtgt accagaagaa tccgatcacc cgtgggctag 60
gagaatgagt gcaacgtttt attgagtgga agtagctctc atctgatggt ggaatcagca 120
gggagatggt tttcccctgg agttgggctt ctaggtggcc ccgtctcttc tctgactgcc 180
ctggccaaac tccaccttgt cctgccggtc aatggcctgc cagcatgcca gtgcctctcg 240
gtgtgctgtt ctgctggcgt gctcctcttg acgaccagct gcttgggtct tccaccgaag 300
tgttcctcac aatgtccagc cgctcgtgtg tctgcccact agggctctcg gtttttatag 360
gcccaggatt ggggcatggc aggccaaggt ggtcttggaa aatggtaaca tttgggcatg 420
aaagcaggag tgcctatcct cacctagatc tgtgggggtg gagccctagc cagggaccca 480
cctttctcta ccagcactt cccctccgc tttccgtatc attttacggg accacactct 540
tcccttccca gcactcccat atcaatagca cataaaacct ttagttctac tgaaaatttg 600
gagtgaatc tctccatagt tagagcacat gtactatttg gtaagcttac ctttggtaat 660
tcatgatacc agctgaagac atcaatacca ataagtgaga agattcttgc caagggtgga 720
aggatttgct tggtaatata gtaagtagca ttcagtctca gagttgggtc ctgcaggact 780
tccactgggc gccttacaag ctggataagt ggtactccgg ggggccata aatgatgacg 840
tatggcactc gctcccaac ctgaggctca gagcgccggt cataagtcag cattttccta 900
ttcgaattca aagaaaaaaa ccacacccaa atactgttat ggaagccatc tttctcacia 960
attaagtcag aattaaaca taataaaaca atgaaaaaac tacaaaagac aatttcacaaa 1020
cagaattact atttctttga gaataatata tatacathtt gcttagttga caccaagtgt 1080
agataccta caagagtta aaaattactc tgtaaatcac tgttacctgg aaaataatga 1140
gttggctgat atcaactagc ttaagatgta ttttaggtgg gagtggtggc tcgtgcctgt 1200
aattccagca ctttgggagg ccaaggatca cctgagccca ggagtttgag accagcttgg 1260
gcgacatagc gagaccccca tctccacaaa aaaattagcc accataatgg tgcacgcctg 1320
ttgtcccagc tgcttgggag gctgatgtgg gaggttcgct tgggcctggg aggtcggggc 1380
tgcagtgagc tgtgattgag cactgcttt ccagtctggg cgacagagtg agaccctgcc 1440
tccaggggca ggaaaaagat gtattttatc ccctaagaat gagaactaac tatcattaaa 1500
tctatcatga aagataagag aaaatttacc tgggttaaact atctgattcc agaattgattc 1560
catctaaatg ctttggtaga acaatattaa ggactaccta gtgattttga tgtgggaaaa 1620
acaaatgaat caaggactgt gtcacaattt ctttgcataa gtttctgaag gaaagcccaa 1680

aagaggattt cctcttcaca tcttacgaag tatgagggaa tattcactaa ctttcagtga 1740
ggcttcatgt gcgcaggtaa agaaacaatt ttgtgtctgt gtctatggac tgtcttttca 1800
tttgggtctt ctagacctgt gcagaatctc aggactatcc tggcaaggta ctgagggtat 1860
cattaccttg taagttcaag ggctggcaca caagctcctg gtttataaga aaaacttcct 1920
ctgtattcct tggcaaagat aaagtcttgt atgctggcct ttccttccag aagcttcata 1980
cattgtcgct gaacatactg ttttaattaga cttatatctc tcgtttcaaa tagcagcttt 2040
agagaacgct caagtatctt aaaacaaaaa agatatgcaa gtaaagataa aatatctcct 2100
agtatgtttc accattgtta ttacactttc aattatgtaa ctactgggta acctgctaac 2160
ctacctaagt aatctgttac ttacaatgat gggataataa aattattaaa aattgaaagt 2220
tatttgagac tttagagagc ttgctatttc taaaaataga atatcatgta gtttaaggaa 2280
aaccatgtgg aatgacatgc aaattaacat ttgttatgtg ataagcaaca taaaacacgt 2340
taagtttttt ttttttaata gtcgaagcct ctaggtacaa gaaaccagac acgatttcat 2400
gtattgtttt tttatttacc agtatagaaa gggctcctgat tcagtcacag gaaacacaga 2460
gaagtttaaa gatcaccagg cagtcaaadc aaaatcagcc tttttgactt tttataaaaa 2520
tgttgataaa taactatgac agaattttct gtgttggaag tttctactag gaaataacca 2580
gcatgcagtg gatttatgta agattaattt tttttttttt tttgagacgg agtcttgctc 2640
tgtcgctcag gctgggggtgc ggtggcgtga tctcggtca ctgcagcctc tgctcccggg 2700
gttcaagcga ctctcctgcc tcagcctccc gagtagctgg gactacaggc gtgtgccacc 2760
acgccccggt aattttttgt tttagtggag atgggggttt gccgtgttg cgggatggt 2820
ctcgatctcc tgacctcatg atccgcccgc ctcggcctcc caaagtgctg ggattatagg 2880
cgtgagccac cgtgccccgc caagtagatt aagttagttc taggtgaaag ctctgtgtct 2940
tattctgaat actggaacca atggatatat atacaataag cttattaatt tagcttgagt 3000
acaattctgc ctacagagaa agaaggtact ggattatatc tcttcagaa atagtgtga 3060
aaaaaatccc aaattcacat aatattatat atatgacata atatattgat aatattcaat 3120
ccactatgct tttaatgctc tcatcaaaca gtatatagaa tgctattccc atttatacca 3180
tttatttttg ttttgctata gaaggagtct aggccccgaat aaagaaaaag ccaaaaaata 3240
tcagggttat tatggttgag aaatttgagt aatttatagc taattgtttg gaaatcttcc 3300
taatttgtat ttgataatcc aaataaatga aagttcagtt aatctggaat ggattatggt 3360
tctctctttt tattttgatg gagtctcact ctcttgccca ggctgagtgc agtgttgtga 3420

tcttggctgt ctgcagcctc ctctctcccg gttcaagtga ttctcacacc tcagtctccc 3480
 gtagtagctgg gattacaggc atgcgccacc acacccggat aatttttgta tttctgggtg 3540
 agatgggggtt tcactacgtt ggccaggctg gtctcgaact cctgatctca agtgatctgc 3600
 ctgccttggc ctcccaaagt gctgggatta caggcgtgag ccaccacacc gggctctgttt 3660
 gtcttgttta tcctgctctg ctttcaggag ttgttacact atgacatata tgatctccag 3720
 gcaataggca ggaaactctc atgtattgct ttttagtaat ttttgataaa acgtaaggta 3780
 tgggaggaag tcagaaactt aaggactaaa atcaaataag ttgtttttat tagaaataaa 3840
 attctagcta ctcgaggaggc tgaggcagaa ggattgcttg agctcgggag gtcagggtgc 3900
 cggtagacca ggattggact actgcactcc agcctgggtg acaaagtgc accctgtctc 3960
 aaacaaacaa aaacaaaata agacactaca aataaaattc tcaaaattca tgtagttact 4020
 cctatagtta tattaattat ataggctata gttccttcac ctagatgcat gtttaagggt 4080
 agttacattt tcaccagaca ctttagccac ttatttctta cagatgaaga caagatagat 4140
 aaaccttacc ttagaaacag cagggcagga atctcttctg actgtttcta ttccttttgc 4200
 atcaaatact gggtccttct gatccagtgt ttcatacatg taaccacat acctcttttt 4260
 tgtttgtaaa acacaggga aatatacctg tggtaaaatt aataaaatgc ctctt . 4315

<210> 40

<211> 685

<212> PRT

<213> Homo sapiens

<400> 40

Met	Val	Leu	Val	Leu	His	His	Ile	Leu	Ile	Ala	Val	Val	Gln	Phe	Leu
1				5				10					15		
Arg	Arg	Gly	Gln	Gln	Val	Phe	Leu	Lys	Pro	Asp	Glu	Pro	Pro	Pro	Pro
			20					25					30		
Pro	Gln	Pro	Cys	Ala	Asp	Ser	Leu	Gln	Asp	Ala	Leu	Leu	Ser	Leu	Gly
			35					40					45		

Ser Val Ile Asp Ile Ser Gly Leu Gln Arg Ala Val Lys Glu Ala Leu
 50 55 60
 Ser Ala Val Leu Pro Arg Val Glu Thr Val Tyr Thr Tyr Leu Leu Asp
 65 70 75 80
 Gly Glu Ser Gln Leu Val Cys Glu Asp Pro Pro His Glu Leu Pro Gln
 85 90 95
 Glu Gly Lys Val Arg Phe Thr Asp Glu Asp Glu His Val Ile Gln His
 100 105 110
 Cys Phe His Tyr Thr Ser Thr Val Leu Thr Ser Thr Leu Ala Phe Gln
 115 120 125
 Lys Glu Gln Lys Leu Lys Cys Glu Cys Gln Ala Leu Leu Gln Val Ala
 130 135 140
 Lys Asn Leu Phe Thr His Leu Asp Asp Val Ser Val Leu Leu Gln Glu
 145 150 155 160
 Ile Ile Thr Glu Ala Arg Asn Leu Ser Asn Ala Glu Ile Cys Ser Val
 165 170 175
 Phe Leu Leu Asp Gln Asn Glu Leu Val Ala Lys Val Phe Asp Gly Gly
 180 185 190
 Val Val Asp Asp Glu Ser Tyr Glu Ile Arg Ile Pro Ala Asp Gln Gly
 195 200 205
 Ile Ala Gly His Val Ala Thr Thr Gly Gln Ile Leu Asn Ile Pro Asp
 210 215 220
 Ala Tyr Ala His Pro Leu Phe Tyr Arg Gly Val Asp Asp Ser Thr Gly
 225 230 235 240
 Phe Arg Thr Arg Asn Ile Leu Cys Phe Pro Ile Lys Asn Glu Asn Gln
 245 250 255
 Glu Val Ile Gly Val Ala Glu Leu Val Asn Lys Ile Asn Gly Pro Trp
 260 265 270
 Phe Ser Lys Phe Asp Glu Asp Leu Ala Thr Ala Phe Ser Ile Tyr Cys

275	280	285
Gly Ile Ser Ile Ala His Ser Leu Leu Tyr Lys Lys Val Asn Glu Ala		
290	295	300
Gln Tyr Arg Ser His Leu Ala Asn Glu Met Met Met Tyr His Met Lys		
305	310	315
Val Ser Asp Asp Glu Tyr Thr Lys Leu Leu His Asp Gly Ile Gln Pro		
325	330	335
Val Ala Ala Ile Asp Ser Asn Phe Ala Ser Phe Thr Tyr Thr Pro Arg		
340	345	350
Ser Leu Pro Glu Asp Asp Thr Ser Met Ala Ile Leu Ser Met Leu Gln		
355	360	365
Asp Met Asn Phe Ile Asn Asn Tyr Lys Ile Asp Cys Pro Thr Leu Ala		
370	375	380
Arg Phe Cys Leu Met Val Lys Lys Gly Tyr Arg Asp Pro Pro Tyr His		
385	390	395
Asn Trp Met His Ala Phe Ser Val Ser His Phe Cys Tyr Leu Leu Tyr		
405	410	415
Lys Asn Leu Glu Leu Thr Asn Tyr Leu Glu Asp Ile Glu Ile Phe Ala		
420	425	430
Leu Phe Ile Ser Cys Met Cys His Asp Leu Asp His Arg Gly Thr Asn		
435	440	445
Asn Ser Phe Gln Val Ala Ser Lys Ser Val Leu Ala Ala Leu Tyr Ser		
450	455	460
Ser Glu Gly Ser Val Met Glu Arg His His Phe Ala Gln Ala Ile Ala		
465	470	475
Ile Leu Asn Thr His Gly Cys Asn Ile Phe Asp His Phe Ser Arg Lys		
485	490	495
Asp Tyr Gln Arg Met Leu Asp Leu Met Arg Asp Ile Ile Leu Ala Thr		
500	505	510

Asp Leu Ala His His Leu Arg Ile Phe Lys Asp Leu Gln Lys Met Ala
515 520 525

Glu Val Gly Tyr Asp Arg Asn Asn Lys Gln His His Arg Leu Leu Leu
530 535 540

Cys Leu Leu Met Thr Ser Cys Asp Leu Ser Asp Gln Thr Lys Gly Trp
545 550 555 560

Lys Thr Thr Arg Lys Ile Ala Glu Leu Ile Tyr Lys Glu Phe Phe Ser
565 570 575

Gln Gly Asp Leu Glu Lys Ala Met Gly Asn Arg Pro Met Glu Met Met
580 585 590

Asp Arg Glu Lys Ala Tyr Ile Pro Glu Leu Gln Ile Ser Phe Met Glu
595 600 605

His Ile Ala Met Pro Ile Tyr Lys Leu Leu Gln Asp Leu Phe Pro Lys
610 615 620

Ala Ala Glu Leu Tyr Glu Arg Val Ala Ser Asn Arg Glu His Trp Thr
625 630 635 640

Lys Val Ser His Lys Phe Thr Ile Arg Gly Leu Pro Ser Asn Asn Ser
645 650 655

Leu Asp Phe Leu Asp Glu Glu Tyr Glu Val Pro Asp Leu Asp Gly Thr
660 665 670

Arg Ala Pro Ile Asn Gly Cys Cys Ser Leu Asp Ala Glu
675 680 685

<210> 41

<211> 563

<212> PRT

<213> Homo sapiens

<400> 41

Met Gly Pro Leu Thr Phe Met Asp Val Ala Ile Glu Phe Ser Leu Glu

1 5 10 15

Glu Trp Gln Cys Leu Asp Thr Ala Gln Gln Asn Leu Tyr Arg Asn Val

20 25 30

Met Leu Glu Asn Tyr Arg Asn Leu Val Phe Leu Gly Ile Ala Val Ser

35 40 45

Asn Leu Asp Leu Ile Thr Cys Leu Glu Gln Gly Lys Glu Pro Trp Asn

50 55 60

Met Lys Arg His Glu Met Val Ala Lys Pro Pro Ala Met Cys Ser His

65 70 75 80

Phe Ala Lys Asp Leu Arg Pro Glu Gln Tyr Ile Lys Asn Ser Phe Gln

85 90 95

Gln Val Ile Leu Arg Arg Tyr Gly Lys Cys Gly Tyr Gln Lys Gly Cys

100 105 110

Lys Ser Val Asp Glu His Lys Leu His Lys Gly Gly His Lys Gly Leu

115 120 125

Asn Gln Cys Val Thr Thr Thr Gln Ser Lys Ile Val Gln Cys Asp Lys

130 135 140

Tyr Val Lys Val Phe His Lys Tyr Ser Asn Ala Lys Arg His Lys Ile

145 150 155 160

Arg His Thr Gly Lys Asn Pro Phe Lys Cys Lys Glu Cys Gly Lys Ser

165 170 175

Phe Cys Met Leu Ser Gln Leu Thr Gln His Glu Ile Ile His Thr Gly

180 185 190

Glu Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Lys Lys Ser

195 200 205

Ser Asn Leu Thr Asn His Lys Ile Ile His Thr Gly Glu Lys Pro Tyr

210 215 220

Lys Cys Glu Glu Cys Gly Lys Ala Phe Asn Gln Ser Ser Thr Leu Thr
 225 230 235 240
 Arg His Lys Ile Ile His Thr Gly Glu Lys Leu Tyr Lys Cys Glu Glu
 245 250 255
 Cys Gly Lys Ala Phe Asn Arg Ser Ser Asn Leu Thr Lys His Lys Ile
 260 265 270
 Val His Thr Gly Glu Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala
 275 280 285
 Phe Lys Gln Ser Ser Asn Leu Thr Asn His Lys Lys Ile His Thr Gly
 290 295 300
 Glu Lys Pro Tyr Lys Cys Gly Glu Cys Gly Lys Ala Phe Thr Leu Ser
 305 310 315 320
 Ser His Leu Thr Thr His Lys Arg Ile His Thr Gly Glu Lys Pro Tyr
 325 330 335
 Lys Cys Glu Glu Cys Gly Lys Ala Phe Ser Val Phe Ser Thr Leu Thr
 340 345 350
 Lys His Lys Ile Ile His Thr Glu Glu Lys Pro Tyr Lys Cys Glu Glu
 355 360 365
 Cys Gly Lys Ala Phe Asn Arg Ser Ser His Leu Thr Asn His Lys Val
 370 375 380
 Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala
 385 390 395 400
 Phe Thr Lys Ser Ser Thr Leu Thr Tyr His Lys Val Ile His Thr Gly
 405 410 415
 Lys Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Ser Ile Phe
 420 425 430
 Ser Ile Leu Thr Lys His Lys Val Ile His Thr Glu Asp Lys Pro Tyr
 435 440 445
 Lys Cys Glu Glu Cys Gly Lys Thr Phe Asn Tyr Ser Ser Asn Phe Thr

450 455 460
 Asn His Lys Lys Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Glu Glu
 465 470 475 480
 Cys Gly Lys Ser Phe Ile Leu Ser Ser His Leu Thr Thr His Lys Ile
 485 490 495
 Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala
 500 505 510
 Phe Asn Gln Ser Ser Thr Leu Met Lys His Lys Ile Ile His Thr Gly
 515 520 525
 Glu Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Asn Gln Ser
 530 535 540
 Pro Asn Leu Thr Lys His Lys Arg Ile His Thr Lys Glu Lys Pro Tyr
 545 550 555 560
 Lys Cys Lys

<210> 42

<211> 1586

<212> PRT

<213> Homo sapiens

<400> 42

Met Glu Glu Met Arg Asp Ser Tyr Leu Glu Glu Asp Val Tyr Gln Leu
 1 5 10 15
 Gln Glu Leu Arg Arg Glu Leu Asp Arg Ala Asn Lys Asn Cys Arg Ile
 20 25 30
 Leu Gln Tyr Arg Leu Arg Lys Ala Glu Gln Lys Ser Leu Lys Val Ala
 35 40 45

Glu Thr Gly Gln Val Asp Gly Glu Leu Ile Arg Ser Leu Glu Gln Asp
 50 55 60
 Leu Lys Val Ala Lys Asp Val Ser Val Arg Leu His His Glu Leu Lys
 65 70 75 80
 Thr Val Glu Glu Lys Arg Ala Lys Ala Glu Asp Glu Asn Glu Thr Leu
 85 90 95
 Arg Gln Gln Met Ile Glu Val Glu Ile Ser Lys Gln Ala Leu Gln Asn
 100 105 110
 Glu Leu Glu Arg Leu Lys Glu Ser Ser Leu Lys Arg Arg Ser Thr Arg
 115 120 125
 Glu Met Tyr Lys Glu Lys Lys Thr Phe Asn Gln Asp Asp Ser Ala Asp
 130 135 140
 Leu Arg Cys Gln Leu Gln Phe Ala Lys Glu Glu Ala Phe Leu Met Arg
 145 150 155 160
 Lys Lys Met Ala Lys Leu Gly Arg Glu Lys Asp Glu Leu Glu Gln Glu
 165 170 175
 Leu Gln Lys Tyr Lys Ser Leu Tyr Gly Asp Val Asp Ser Pro Leu Pro
 180 185 190
 Thr Gly Glu Ala Gly Gly Pro Pro Ser Thr Arg Glu Ala Glu Leu Lys
 195 200 205
 Leu Arg Leu Lys Leu Val Glu Glu Glu Ala Asn Ile Leu Gly Arg Lys
 210 215 220
 Ile Val Glu Leu Glu Val Glu Asn Arg Gly Leu Lys Ala Glu Met Glu
 225 230 235 240
 Asp Thr Arg Gly Gln Gln Glu Arg Glu Gly Pro Gly Arg Asp His Ala
 245 250 255
 Pro Ser Ile Pro Thr Ser Pro Phe Gly Asp Ser Leu Glu Ser Ser Thr
 260 265 270
 Glu Leu Arg Arg His Leu Gln Phe Val Glu Glu Glu Ala Glu Leu Leu

275	280	285
Arg Arg Ser Ile Ser Glu Ile Glu Asp His Asn Arg Gln Leu Thr His		
290	295	300
Glu Leu Ser Lys Phe Lys Phe Glu Pro Pro Arg Glu Pro Gly Trp Leu		
305	310	315
Gly Glu Gly Ala Ser Pro Gly Ala Gly Gly Gly Ala Pro Leu Gln Glu		
325	330	335
Glu Leu Lys Ser Ala Arg Leu Gln Ile Ser Glu Leu Ser Gly Lys Val		
340	345	350
Leu Lys Leu Gln His Glu Asn His Ala Leu Leu Ser Asn Ile Gln Arg		
355	360	365
Cys Asp Leu Ala Ala His Leu Gly Leu Arg Ala Pro Ser Pro Arg Asp		
370	375	380
Ser Asp Ala Glu Ser Asp Ala Gly Lys Lys Glu Ser Asp Gly Glu Glu		
385	390	395
Ser Arg Leu Pro Gln Pro Lys Arg Glu Gly Pro Val Gly Gly Glu Ser		
405	410	415
Asp Ser Glu Glu Met Phe Glu Lys Thr Ser Gly Phe Gly Ser Gly Lys		
420	425	430
Pro Ser Glu Ala Ser Glu Pro Cys Pro Thr Glu Leu Leu Lys Ala Arg		
435	440	445
Glu Asp Ser Glu Tyr Leu Val Thr Leu Lys His Glu Ala Gln Arg Leu		
450	455	460
Glu Arg Thr Val Glu Arg Leu Ile Thr Asp Thr Asp Ser Phe Leu His		
465	470	475
Asp Ala Gly Leu Arg Gly Gly Ala Pro Leu Pro Gly Pro Gly Leu Gln		
485	490	495
Gly Glu Glu Glu Gln Gly Glu Gly Asp Gln Gln Glu Pro Gln Leu Leu		
500	505	510

Gly Thr Ile Asn Ala Lys Met Lys Ala Phe Lys Lys Glu Leu Gln Ala
515 520 525
Phe Leu Glu Gln Val Asn Arg Ile Gly Gly Gly Leu Ser Pro Leu Pro
530 535 540
His Leu Thr Glu Ser Ser Ser Phe Leu Ser Thr Val Thr Ser Val Ser
545 550 555 560
Arg Asp Ser Pro Ile Gly Asn Leu Gly Lys Glu Leu Gly Pro Asp Leu
565 570 575
Gln Ser Arg Leu Lys Glu Gln Leu Glu Trp Gln Leu Gly Pro Ala Arg
580 585 590
Gly Asp Glu Arg Glu Ser Leu Arg Leu Arg Ala Ala Arg Glu Leu His
595 600 605
Arg Arg Ala Asp Gly Asp Thr Gly Ser His Gly Leu Gly Gly Gln Thr
610 615 620
Cys Phe Ser Leu Glu Leu Arg Gly Pro Pro Val Leu Pro Glu Gln Ser
625 630 635 640
Val Ser Ile Glu Glu Leu Gln Gly Gln Leu Val Gln Ala Ala Arg Leu
645 650 655
His Gln Glu Glu Thr Glu Thr Phe Thr Asn Lys Ile His Lys Met Glu
660 665 670
Glu Glu His Leu Tyr Ala Leu Arg Trp Lys Glu Leu Glu Met His Ser
675 680 685
Leu Ala Leu Gln Asn Thr Leu His Glu Arg Thr Trp Ser Asp Glu Lys
690 695 700
Asn Leu Met Gln Gln Glu Leu Arg Ser Leu Lys Gln Asn Ile Phe Leu
705 710 715 720
Phe Tyr Val Lys Leu Arg Trp Leu Leu Lys His Trp Arg Gln Gly Lys
725 730 735
Gln Met Glu Glu Glu Gly Glu Glu Phe Thr Glu Gly Glu His Pro Glu

740	745	750
Thr Leu Ser Arg Leu Gly Glu Leu Gly Val Gln Gly Gly His Gln Ala		
755	760	765
Asp Gly Pro Asp His Asp Ser Asp Arg Gly Cys Gly Phe Pro Val Gly		
770	775	780
Glu His Ser Pro His Ser Arg Val Gln Ile Gly Asp His Ser Leu Arg		
785	790	795
800		
Leu Gln Thr Ala Asp Arg Gly Gln Pro His Lys Gln Val Val Glu Asn		
805	810	815
Gln Gln Leu Phe Ser Ala Phe Lys Ala Leu Leu Glu Asp Phe Arg Ala		
820	825	830
Glu Leu Arg Glu Asp Glu Arg Ala Arg Leu Arg Leu Gln Gln Gln Tyr		
835	840	845
Ala Ser Asp Lys Ala Ala Trp Asp Val Glu Trp Ala Val Leu Lys Cys		
850	855	860
Arg Leu Glu Gln Leu Glu Glu Lys Thr Glu Asn Lys Leu Gly Glu Leu		
865	870	875
880		
Gly Ser Ser Ala Glu Ser Lys Gly Ala Leu Lys Lys Glu Arg Glu Val		
885	890	895
His Gln Lys Leu Leu Ala Asp Ser His Ser Leu Val Met Asp Leu Arg		
900	905	910
Trp Gln Ile His His Ser Glu Lys Asn Trp Asn Arg Glu Lys Val Glu		
915	920	925
Leu Leu Asp Arg Leu Asp Arg Asp Arg Gln Glu Trp Glu Arg Gln Lys		
930	935	940
Lys Glu Phe Leu Trp Arg Ile Glu Gln Leu Gln Lys Glu Asn Ser Pro		
945	950	955
960		
Arg Arg Gly Gly Ser Phe Leu Cys Asp Gln Lys Asp Gly Asn Val Arg		
965	970	975

Pro Phe Pro His Gln Gly Ser Leu Arg Met Pro Arg Pro Val Ala Met
980 985 990
Trp Pro Cys Ala Asp Ala Asp Ser Ile Pro Phe Glu Asp Arg Pro Leu
995 1000 1005
Ser Lys Leu Lys Glu Ser Asp Arg Cys Ser Ala Ser Glu Asn Leu Tyr
1010 1015 1020
Leu Asp Ala Leu Ser Leu Asp Asp Glu Pro Glu Glu Pro Pro Ala His
1025 1030 1035 1040
Arg Pro Glu Arg Glu Phe Arg Asn Arg Leu Pro Glu Glu Glu Glu Asn
1045 1050 1055
His Lys Gly Asn Leu Gln Arg Ala Val Ser Val Ser Ser Met Ser Glu
1060 1065 1070
Phe Gln Arg Leu Met Asp Ile Ser Pro Phe Leu Pro Glu Lys Gly Leu
1075 1080 1085
Pro Ser Thr Ser Ser Lys Glu Asp Val Thr Pro Pro Leu Ser Pro Asp
1090 1095 1100
Asp Leu Lys Tyr Ile Glu Glu Phe Asn Lys Ser Trp Asp Tyr Thr Pro
1105 1110 1115 1120
Asn Arg Gly His Asn Gly Gly Gly Pro Asp Leu Trp Ala Asp Arg Thr
1125 1130 1135
Glu Val Gly Arg Ala Gly His Glu Asp Ser Thr Glu Pro Phe Pro Asp
1140 1145 1150
Ser Ser Trp Tyr Leu Thr Thr Ser Val Thr Met Thr Thr Asp Thr Met
1155 1160 1165
Thr Ser Pro Glu His Cys Gln Lys Gln Pro Leu Arg Ser His Val Leu
1170 1175 1180
Thr Glu Gln Ser Gly Leu Arg Val Leu His Ser Pro Pro Ala Val Arg
1185 1190 1195 1200
Arg Val Asp Ser Ile Thr Ala Ala Gly Gly Glu Gly Pro Phe Pro Thr

1205	1210	1215
Ser Arg Ala Arg Gly Ser Pro Gly Asp Thr Lys Gly Gly Pro Pro Glu		
1220	1225	1230
Pro Met Leu Ser Arg Trp Pro Cys Thr Ser Pro Arg His Ser Arg Asp		
1235	1240	1245
Tyr Val Glu Gly Ala Arg Arg Pro Leu Asp Ser Pro Leu Cys Thr Ser		
1250	1255	1260
Leu Gly Phe Ala Ser Pro Leu His Ser Leu Glu Met Ser Lys Asn Leu		
1265	1270	1275
Ser Asp Asp Met Lys Glu Val Ala Phe Ser Val Arg Asn Ala Ile Cys		
1285	1290	1295
Ser Gly Pro Gly Glu Leu Gln Val Lys Asp Met Ala Cys Gln Thr Asn		
1300	1305	1310
Gly Ser Arg Thr Met Gly Thr Gln Thr Val Gln Thr Ile Ser Val Gly		
1315	1320	1325
Leu Gln Thr Glu Ala Leu Arg Gly Ser Gly Val Thr Ser Ser Pro His		
1330	1335	1340
Lys Cys Leu Thr Pro Lys Ala Gly Gly Gly Ala Thr Pro Val Ser Ser		
1345	1350	1355
Pro Ser Arg Ser Leu Arg Ser Arg Gln Val Ala Pro Ala Ile Glu Lys		
1365	1370	1375
Val Gln Ala Lys Phe Glu Arg Thr Cys Cys Ser Pro Lys Tyr Gly Ser		
1380	1385	1390
Pro Lys Leu Gln Arg Lys Pro Leu Pro Lys Ala Asp Gln Pro Asn Asn		
1395	1400	1405
Arg Thr Ser Pro Gly Met Ala Gln Lys Gly Tyr Ser Glu Ser Ala Trp		
1410	1415	1420
Ala Arg Ser Thr Thr Thr Arg Glu Ser Pro Val His Thr Thr Ile Asn		
1425	1430	1435
		1440

Asp Gly Leu Ser Ser Leu Phe Asn Ile Ile Asp His Ser Pro Val Val

1445

1450

1455

Gln Asp Pro Phe Gln Lys Gly Leu Arg Ala Gly Ser Arg Ser Arg Ser

1460

1465

1470

Ala Glu Pro Arg Pro Glu Leu Gly Pro Gly Gln Glu Thr Gly Thr Asn

1475

1480

1485

Ser Arg Gly Arg Ser Pro Ser Pro Ile Gly Val Gly Ser Glu Met Cys

1490

1495

1500

Arg Glu Glu Gly Gly Glu Gly Thr Pro Val Lys Gln Asp Leu Ser Ala

1505

1510

1515

1520

Pro Pro Gly Tyr Thr Leu Thr Glu Asn Val Ala Arg Ile Leu Asn Lys

1525

1530

1535

Lys Leu Leu Glu His Ala Leu Lys Glu Glu Arg Arg Gln Ala Ala His

1540

1545

1550

Gly Pro Pro Gly Leu His Ser Asp Ser His Ser Leu Gly Asp Thr Ala

1555

1560

1565

Glu Pro Gly Pro Met Glu Asn Gln Thr Val Leu Leu Thr Ala Pro Trp

1570

1575

1580

Gly Leu

1585

<210> 43

<211> 1331

<212> PRT

<213> Homo sapiens

<400> 43

Met Glu Ser His Ser Arg Ala Gly Lys Ser Arg Lys Ser Ala Lys Phe

1

5

10

15

Arg Ser Ile Ser Arg Ser Leu Met Leu Cys Asn Ala Lys Thr Ser Asp
 20 25 30
 Asp Gly Ser Ser Pro Asp Glu Lys Tyr Pro Asp Pro Phe Glu Ile Ser
 35 40 45
 Leu Ala Gln Gly Lys Glu Gly Ile Phe His Ser Ser Val Gln Leu Ala
 50 55 60
 Asp Thr Ser Glu Ala Gly Pro Ser Ser Val Pro Asp Leu Ala Leu Ala
 65 70 75 80
 Ser Glu Ala Ala Gln Leu Gln Ala Ala Gly Asn Asp Arg Gly Lys Thr
 85 90 95
 Cys Arg Arg Ile Phe Phe Met Lys Glu Ser Ser Thr Ala Ser Ser Arg
 100 105 110
 Glu Lys Pro Gly Lys Leu Glu Ala Gln Ser Ser Asn Phe Leu Phe Pro
 115 120 125
 Lys Ala Cys His Gln Arg Ala Arg Ser Asn Ser Thr Ser Val Asn Pro
 130 135 140
 Tyr Cys Thr Arg Glu Ile Asp Phe Pro Met Thr Lys Lys Ser Ala Ala
 145 150 155 160
 Pro Thr Asp Arg Gln Pro Tyr Ser Leu Cys Ser Asn Arg Lys Ser Leu
 165 170 175
 Ser Gln Gln Leu Asp Cys Pro Ala Gly Lys Ala Ala Gly Thr Ser Arg
 180 185 190
 Pro Thr Arg Ser Leu Ser Thr Ala Gln Leu Val Gln Pro Ser Gly Gly
 195 200 205
 Leu Gln Ala Ser Val Ile Ser Asn Ile Val Leu Met Lys Gly Gln Ala
 210 215 220
 Lys Gly Leu Gly Phe Ser Ile Val Gly Gly Lys Asp Ser Ile Tyr Gly
 225 230 235 240
 Pro Ile Gly Ile Tyr Val Lys Thr Ile Phe Ala Gly Gly Ala Ala Ala

245	250	255
Ala Asp Gly Arg Leu Gln Glu Gly Asp Glu Ile Leu Glu Leu Asn Gly		
260	265	270
Glu Ser Met Ala Gly Leu Thr His Gln Asp Ala Leu Gln Lys Phe Lys		
275	280	285
Gln Ala Lys Lys Gly Leu Leu Thr Leu Thr Val Arg Thr Arg Leu Thr		
290	295	300
Ala Pro Pro Ser Leu Cys Ser His Leu Ser Pro Pro Leu Cys Arg Ser		
305	310	315
Leu Ser Ser Ser Thr Cys Ile Thr Lys Asp Ser Ser Ser Phe Ala Leu		
325	330	335
Glu Ser Pro Ser Ala Pro Ile Ser Thr Ala Lys Pro Asn Tyr Arg Ile		
340	345	350
Met Val Glu Val Ser Leu Gln Lys Glu Ala Gly Val Gly Leu Gly Ile		
355	360	365
Gly Leu Cys Ser Val Pro Tyr Phe Gln Cys Ile Ser Gly Ile Phe Val		
370	375	380
His Thr Leu Ser Pro Gly Ser Val Ala His Leu Asp Gly Arg Leu Arg		
385	390	395
Cys Gly Asp Glu Ile Val Glu Ile Ser Asp Ser Pro Val His Cys Leu		
405	410	415
Thr Leu Asn Glu Val Tyr Thr Ile Leu Ser His Cys Asp Pro Gly Pro		
420	425	430
Val Pro Ile Ile Val Ser Arg His Pro Asp Pro Gln Val Ser Glu Gln		
435	440	445
Gln Leu Lys Glu Ala Val Ala Gln Ala Val Glu Asn Thr Lys Phe Gly		
450	455	460
Lys Glu Arg His Gln Trp Ser Leu Glu Gly Val Lys Arg Leu Glu Ser		
465	470	475
		480

Ser Trp His Gly Arg Pro Thr Leu Glu Lys Glu Arg Glu Lys Asn Ser
 485 490 495
 Ala Pro Pro His Arg Arg Ala Gln Lys Val Met Ile Arg Ser Ser Ser
 500 505 510
 Asp Ser Ser Tyr Met Ser Gly Ser Pro Gly Gly Ser Pro Gly Ser Gly
 515 520 525
 Ser Ala Glu Lys Pro Ser Ser Asp Val Asp Ile Ser Thr His Ser Pro
 530 535 540
 Ser Leu Pro Leu Ala Arg Glu Pro Val Val Leu Ser Ile Ala Ser Ser
 545 550 555 560
 Arg Leu Pro Gln Glu Ser Pro Pro Leu Pro Glu Ser Arg Asp Ser His
 565 570 575
 Pro Pro Leu Arg Leu Lys Lys Ser Phe Glu Ile Leu Val Arg Lys Pro
 580 585 590
 Met Ser Ser Lys Pro Lys Pro Pro Pro Arg Lys Tyr Phe Lys Ser Asp
 595 600 605
 Ser Asp Pro Gln Lys Ser Leu Glu Glu Arg Glu Asn Ser Ser Cys Ser
 610 615 620
 Ser Gly His Thr Pro Pro Thr Cys Gly Gln Glu Ala Arg Glu Leu Leu
 625 630 635 640
 Pro Leu Leu Leu Pro Gln Glu Asp Thr Ala Gly Arg Ser Pro Ser Ala
 645 650 655
 Ser Ala Gly Cys Pro Gly Pro Gly Ile Gly Pro Gln Thr Lys Ser Ser
 660 665 670
 Thr Glu Gly Glu Pro Gly Trp Arg Arg Ala Ser Pro Val Thr Gln Thr
 675 680 685
 Ser Pro Ile Lys His Pro Leu Leu Lys Arg Gln Ala Arg Met Asp Tyr
 690 695 700
 Ser Phe Asp Thr Thr Ala Glu Asp Pro Trp Val Arg Ile Ser Asp Cys

705 710 715 720
Ile Lys Asn Leu Phe Ser Pro Ile Met Ser Glu Asn His Gly His Met
 725 730 735
Pro Leu Gln Pro Asn Ala Ser Leu Asn Glu Glu Glu Gly Thr Gln Gly
 740 745 750
His Pro Asp Gly Thr Pro Pro Lys Leu Asp Thr Ala Asn Gly Thr Pro
 755 760 765
Lys Val Tyr Lys Ser Ala Asp Ser Ser Thr Val Lys Lys Gly Pro Pro
 770 775 780
Val Ala Pro Lys Pro Ala Trp Phe Arg Gln Ser Leu Lys Gly Leu Arg
785 790 795 800
Asn Arg Ala Ser Asp Pro Arg Gly Leu Pro Asp Pro Ala Leu Ser Thr
 805 810 815
Gln Pro Ala Pro Ala Ser Arg Glu His Leu Gly Ser His Ile Arg Ala
 820 825 830
Ser Ser Ser Ser Ser Ser Ile Arg Gln Arg Ile Ser Ser Phe Glu Thr
 835 840 845
Phe Gly Ser Ser Gln Leu Pro Asp Lys Gly Ala Gln Arg Leu Ser Leu
 850 855 860
Gln Pro Ser Ser Gly Glu Ala Ala Lys Pro Leu Gly Lys His Glu Glu
865 870 875 880
Gly Arg Phe Ser Gly Leu Leu Gly Arg Gly Ala Ala Pro Thr Leu Val
 885 890 895
Pro Gln Gln Pro Glu Gln Val Leu Ser Ser Gly Ser Pro Ala Ala Ser
 900 905 910
Glu Ala Arg Asp Pro Gly Val Ser Glu Ser Pro Pro Pro Gly Arg Gln
 915 920 925
Pro Asn Gln Lys Thr Leu Pro Pro Gly Pro Asp Pro Leu Leu Arg Leu
 930 935 940

Leu Ser Thr Gln Ala Glu Glu Ser Gln Gly Pro Val Leu Lys Met Pro
 945 950 955 960
 Ser Gln Arg Ala Arg Ser Phe Pro Leu Thr Arg Ser Gln Ser Cys Glu
 965 970 975
 Thr Lys Leu Leu Asp Glu Lys Thr Ser Lys Leu Tyr Ser Ile Ser Ser
 980 985 990
 Gln Val Ser Ser Ala Val Met Lys Ser Leu Leu Cys Leu Pro Ser Ser
 995 1000 1005
 Ile Ser Cys Ala Gln Thr Pro Cys Ile Pro Lys Glu Gly Ala Ser Pro
 1010 1015 1020
 Thr Ser Ser Ser Asn Glu Asp Ser Ala Ala Asn Gly Ser Ala Glu Thr
 1025 1030 1035 1040
 Ser Ala Leu Asp Thr Gly Phe Ser Leu Asn Leu Ser Glu Leu Arg Glu
 1045 1050 1055
 Tyr Thr Glu Gly Leu Thr Glu Ala Lys Glu Asp Asp Asp Gly Asp His
 1060 1065 1070
 Ser Ser Leu Gln Ser Gly Gln Ser Val Ile Ser Leu Leu Ser Ser Glu
 1075 1080 1085
 Glu Leu Lys Lys Leu Ile Glu Glu Val Lys Val Leu Asp Glu Ala Thr
 1090 1095 1100
 Leu Lys Gln Leu Asp Gly Ile His Val Thr Ile Leu His Lys Glu Glu
 1105 1110 1115 1120
 Gly Ala Gly Leu Gly Phe Ser Leu Ala Gly Gly Ala Asp Leu Glu Asn
 1125 1130 1135
 Lys Val Ile Thr Val His Arg Val Phe Pro Asn Gly Leu Ala Ser Gln
 1140 1145 1150
 Glu Gly Thr Ile Gln Lys Gly Asn Glu Val Leu Ser Ile Asn Gly Lys
 1155 1160 1165
 Ser Leu Lys Gly Thr Thr His His Asp Ala Leu Ala Ile Leu Arg Gln

1170 1175 1180
Ala Arg Glu Pro Arg Gln Ala Val Ile Val Thr Arg Lys Leu Thr Pro
1185 1190 1195 1200
Glu Ala Met Pro Asp Leu Asn Ser Ser Thr Asp Ser Ala Ala Ser Ala
1205 1210 1215
Ser Ala Ala Ser Asp Val Ser Val Glu Ser Thr Glu Ala Thr Val Cys
1220 1225 1230
Thr Val Thr Leu Glu Lys Met Ser Ala Gly Leu Gly Phe Ser Leu Glu
1235 1240 1245
Gly Gly Lys Gly Ser Leu His Gly Asp Lys Pro Leu Thr Ile Asn Arg
1250 1255 1260
Ile Phe Lys Gly Ala Ala Ser Glu Gln Ser Glu Thr Val Gln Pro Gly
1265 1270 1275 1280
Asp Glu Ile Leu Gln Leu Gly Gly Thr Ala Met Gln Gly Leu Thr Arg
1285 1290 1295
Phe Glu Ala Trp Asn Ile Ile Lys Ala Leu Pro Asp Gly Pro Val Thr
1300 1305 1310
Ile Val Ile Arg Arg Lys Ser Leu Gln Ser Lys Glu Thr Thr Ala Ala
1315 1320 1325
Gly Asp Ser
1330

<210> 44

<211> 264

<212> PRT

<213> Homo sapiens

<400> 44

Met Lys Ala Gln Leu Asn Pro His Leu Asn Leu Leu Thr Pro Ala Arg
 1 5 10 15
 Leu Pro Ser Gly His Ala Pro Gly Gln Asp Arg Ile Ala Leu Glu Pro
 20 25 30
 Pro Arg Thr Ser Pro Asn Leu Leu Leu Ala Gly Glu Ala Pro Ser Thr
 35 40 45
 Arg Pro Gln Arg Glu Ala Glu Gly Glu Pro Thr Ala Arg Thr Ser Thr
 50 55 60
 Ser Gly Ser Ser Asp Leu Pro Lys Ala Pro Gly Gly His Leu His Leu
 65 70 75 80
 Gln Ala Arg Arg Ala Gly Gln Asn Ser Asp Ala Gly Pro Pro Arg Gly
 85 90 95
 Leu Thr Trp Val Pro Glu Gly Ala Gly Pro Val Leu Gly Gly Cys Gly
 100 105 110
 Arg Gly Met Asp Gln Ser Cys Leu Ser Ala Asp Gly Ala Gly Arg Gly
 115 120 125
 Cys Gly Arg Ala Thr Trp Ser Val Arg Glu Glu Gln Val Lys Gln Trp
 130 135 140
 Ala Ala Glu Met Leu Val Ala Leu Glu Ala Leu His Glu Gln Gly Val
 145 150 155 160
 Leu Cys Arg Asp Leu His Pro Gly Asn Leu Leu Leu Asp Gln Ala Gly
 165 170 175
 Arg Cys Pro Pro His Pro Ser Val Gly Trp Pro Pro Lys Arg Pro Gly
 180 185 190
 Arg Gly Arg Glu Ile Pro Lys Val Glu Gly Thr Ser Lys Val Val Cys
 195 200 205
 Ser Ser Leu Ala Pro Ala Ser Leu Ser Gly Gly Leu Gly Trp Ser Leu
 210 215 220
 Leu Ala Glu Ala Pro Gly Ser Thr Leu Trp Ala Leu Trp Ala Ser Glu

225 230 235 240
Cys Cys Ala Ser Val Glu Gly Ala Val Pro Leu Pro Ser Phe Cys Pro
 245 250 255
Ala Arg Pro Leu Ala Pro Thr Ala
 260

<210> 45

<211> 213

<212> PRT

<213> Homo sapiens

<400> 45

Met Gly Asp Gly Lys Ala Pro Leu His Thr Pro Lys Ala Trp Val Ser
1 5 10 15
Pro Ser Ile Pro Val Leu Val Pro Gly Ile Ile Val Asn Lys Ala Pro
 20 25 30
Leu Glu Lys Val Pro Asp Leu Ile Ser Gln Val Leu Ala Thr Tyr Pro
 35 40 45
Ala Asp Gly Glu Met Ala Glu Ala Ser Cys Gly Val Phe Trp Leu Leu
 50 55 60
Ser Leu Leu Gly Cys Ile Lys Glu Gln Gln Phe Glu Gln Val Val Ala
65 70 75 80
Leu Leu Leu Gln Ser Ile Arg Leu Cys Gln Asp Arg Ala Leu Leu Val
 85 90 95
Asn Asn Ala Tyr Arg Gly Leu Ala Ser Leu Val Lys Val Ser Glu Leu
 100 105 110
Ala Ala Phe Lys Val Val Val Gln Glu Glu Gly Gly Ser Gly Leu Ser
 115 120 125

Leu Ile Lys Glu Thr Tyr Gln Leu His Arg Asp Asp Pro Glu Val Val
 130 135 140
 Glu Asn Val Gly Met Leu Leu Val His Leu Ala Ser Tyr Glu Glu Ile
 145 150 155 160
 Leu Pro Glu Leu Val Ser Ser Ser Met Lys Ala Leu Leu Gln Glu Ile
 165 170 175
 Lys Glu Arg Phe Thr Ser Ser Leu Val Ser Asp Ser Ser Ala Phe Ser
 180 185 190
 Lys Pro Gly Leu Pro Pro Gly Gly Ser Pro Gln Leu Gly Cys Thr Thr
 195 200 205
 Ser Gly Gly Leu Glu
 210

<210> 46

<211> 796

<212> PRT

<213> Homo sapiens

<400> 46

Met Pro Ala Trp Glu Pro Gly Val Ala Ala Ser Cys Gly Trp Arg Ala
 1 5 10 15
 Pro Pro Cys Ser Pro Leu Arg Leu Cys Ile Ala Pro His Pro Cys Thr
 20 25 30
 Pro Pro Ser Ser His Gln Gly Asn Leu Thr Cys Arg Leu Pro Pro Arg
 35 40 45
 Ser Phe Gly Phe Cys Arg Cys Val Trp Val Arg Thr Arg Gly Pro Ser
 50 55 60
 Leu Pro Gly Glu Gln Val Ser Leu Ala Ala His Glu Ser Glu Gly Ala

出証特 2 0 0 4 - 3 0 5 9 6 6 3

His Ala Met His Arg Asp Leu Cys Pro Gly Arg Val Gly Leu Cys Pro
305 310 315 320
Arg Met Asp Pro Val Asp Gly Thr Gln Leu Leu Lys Tyr Ile Arg Asn
325 330 335
Val Asn Phe Ser Gly Ile Ala Gly Asn Pro Val Thr Phe Asn Glu Asn
340 345 350
Gly Asp Ala Pro Gly Arg Tyr Asp Ile Tyr Gln Tyr Gln Leu Arg Asn
355 360 365
Asp Ser Ala Glu Tyr Lys Val Ile Gly Ser Trp Thr Asp His Leu His
370 375 380
Leu Arg Ile Glu Arg Met His Trp Pro Gly Ser Gly Gln Gln Leu Pro
385 390 395 400
Arg Ser Ile Cys Ser Leu Pro Cys Gln Pro Gly Glu Arg Lys Lys Thr
405 410 415
Val Lys Gly Met Pro Cys Cys Trp His Cys Glu Pro Cys Thr Gly Tyr
420 425 430
Gln Tyr Gln Val Asp Arg Tyr Thr Cys Lys Thr Cys Pro Tyr Asp Met
435 440 445
Arg Pro Thr Glu Asn Arg Thr Gly Cys Arg Pro Ile Pro Ile Ile Lys
450 455 460
Leu Glu Trp Gly Ser Pro Trp Ala Val Leu Pro Leu Phe Leu Ala Val
465 470 475 480
Val Gly Ile Ala Ala Thr Leu Phe Val Val Ile Thr Phe Val Arg Tyr
485 490 495
Asn Asp Thr Pro Ile Val Lys Ala Ser Gly Arg Glu Leu Ser Tyr Val
500 505 510
Leu Leu Ala Gly Ile Phe Leu Cys Tyr Ala Thr Thr Phe Leu Met Ile
515 520 525
Ala Glu Pro Asp Leu Gly Thr Cys Ser Leu Arg Arg Ile Phe Leu Gly

530 535 540
Leu Gly Met Ser Ile Ser Tyr Ala Ala Leu Leu Thr Lys Thr Asn Arg
545 550 555 560
Ile Tyr Arg Ile Phe Glu Gln Gly Lys Arg Ser Val Ser Ala Pro Arg
 565 570 575
Phe Ile Ser Pro Ala Ser Gln Leu Ala Ile Thr Phe Ser Leu Ile Ser
 580 585 590
Leu Gln Leu Leu Gly Ile Cys Val Trp Phe Val Val Asp Pro Ser His
 595 600 605
Ser Val Val Asp Phe Gln Asp Gln Arg Thr Leu Asp Pro Arg Phe Ala
 610 615 620
Arg Gly Val Leu Lys Cys Asp Ile Ser Asp Leu Ser Leu Ile Cys Leu
625 630 635 640
Leu Gly Tyr Ser Met Leu Leu Met Val Thr Cys Thr Val Tyr Ala Ile
 645 650 655
Lys Thr Arg Gly Val Pro Glu Thr Phe Asn Glu Ala Lys Pro Ile Gly
 660 665 670
Phe Thr Met Tyr Thr Thr Cys Ile Val Trp Leu Ala Phe Ile Pro Ile
 675 680 685
Phe Phe Gly Thr Ser Gln Ser Ala Asp Lys Leu Tyr Ile Gln Thr Thr
 690 695 700
Thr Leu Thr Val Ser Val Ser Leu Ser Ala Ser Val Ser Leu Gly Met
705 710 715 720
Leu Tyr Met Pro Lys Val Tyr Ile Ile Leu Phe His Pro Glu Gln Asn
 725 730 735
Val Pro Lys Arg Lys Arg Ser Leu Lys Ala Val Val Thr Ala Ala Thr
 740 745 750
Met Ser Asn Lys Phe Thr Gln Lys Gly Asn Phe Arg Pro Asn Gly Glu
 755 760 765

Ala Lys Ser Glu Leu Cys Glu Asn Leu Glu Ala Pro Ala Leu Ala Thr

770

775

780

Lys Gln Thr Tyr Val Thr Tyr Thr Asn His Ala Ile

785

790

795

<210> 47

<211> 560

<212> PRT

<213> Homo sapiens

<400> 47

Met Glu Lys Gly Thr Glu Leu Leu Val Ser Pro Ser Gln Ser Gly Pro

1

5

10

15

Gly Gly Asp Gln Pro Leu Leu Val Lys His Arg Glu Asp Ser Ala Val

20

25

30

Val Leu Ser Leu Ile Asp Thr Ile Asp Thr Val Met Gly His Val Ser

35

40

45

Ser Asn Leu His Gly Ser Thr Pro Gln Val Thr Val Glu Gly Ser Ser

50

55

60

Ala Met Ala Glu Phe Ser Val Ala Lys Ile Leu Pro Lys Thr Val Asn

65

70

75

80

Ser Ser His Tyr Arg Phe Pro Ala His Gly Gln Ser Phe Ile Gln Ile

85

90

95

Pro His Glu Ala Phe His Arg His Ala Trp Ser Thr Val Val Gly Leu

100

105

110

Leu Tyr His Ser Met His Tyr Tyr Leu Asn Asn Ile Trp Pro Ala His

115

120

125

Thr Lys Ile Ala Glu Ala Met His His Gln Asp Cys Leu Leu Phe Ala

130 135 140
Thr Ser His Leu Ile Ser Leu Glu Val Ser Pro Pro Pro Thr Leu Ser
145 150 155 160
Gln Asn Leu Ser Gly Ser Pro Leu Ile Thr Val His Leu Lys His Arg
165 170 175
Leu Thr Arg Lys Gln His Ser Glu Ala Thr Asn Ser Ser Asn Arg Val
180 185 190
Phe Val Tyr Cys Ala Phe Leu Asp Phe Ser Ser Gly Glu Gly Val Trp
195 200 205
Ser Asn His Gly Cys Ala Leu Thr Arg Gly Asn Leu Thr Tyr Ser Val
210 215 220
Cys Arg Cys Thr His Leu Thr Asn Phe Ala Ile Leu Met Gln Val Val
225 230 235 240
Pro Leu Glu Leu Ala Arg Gly His Gln Val Ala Leu Ser Ser Ile Ser
245 250 255
Tyr Val Gly Cys Ser Leu Ser Val Leu Cys Leu Val Ala Thr Leu Val
260 265 270
Thr Phe Ala Val Leu Ser Ser Val Ser Thr Ile Arg Asn Gln Arg Tyr
275 280 285
His Ile His Ala Asn Leu Ser Phe Ala Val Leu Val Ala Gln Val Leu
290 295 300
Leu Leu Ile Ser Phe Arg Leu Glu Pro Gly Thr Thr Pro Cys Gln Val
305 310 315 320
Met Ala Val Leu Leu His Tyr Phe Phe Leu Ser Ala Phe Ala Trp Met
325 330 335
Leu Val Glu Gly Leu His Leu Tyr Ser Met Val Ile Lys Val Phe Gly
340 345 350
Ser Glu Asp Ser Lys His Arg Tyr Tyr Tyr Gly Met Gly Trp Gly Phe
355 360 365

Pro Leu Leu Ile Cys Ile Ile Ser Leu Ser Phe Ala Met Asp Ser Tyr
370 375 380
Gly Thr Ser Asn Asn Cys Trp Leu Ser Leu Ala Ser Gly Ala Ile Trp
385 390 395 400
Ala Phe Val Ala Pro Ala Leu Phe Val Ile Val Val Asn Ile Gly Ile
405 410 415
Leu Ile Ala Val Thr Arg Val Ile Ser Gln Ile Ser Ala Asp Asn Tyr
420 425 430
Lys Ile His Gly Asp Pro Ser Ala Phe Lys Leu Thr Ala Lys Ala Val
435 440 445
Ala Val Leu Leu Pro Ile Leu Gly Thr Ser Trp Val Phe Gly Val Leu
450 455 460
Ala Val Asn Gly Cys Ala Val Val Phe Gln Tyr Met Phe Ala Thr Leu
465 470 475 480
Asn Ser Leu Gln Gly Leu Phe Ile Phe Leu Phe His Cys Leu Leu Asn
485 490 495
Ser Glu Val Arg Ala Ala Phe Lys His Lys Thr Lys Val Trp Ser Leu
500 505 510
Thr Ser Ser Ser Ala Arg Thr Ser Asn Ala Lys Pro Phe His Ser Asp
515 520 525
Leu Met Asn Gly Thr Arg Pro Gly Met Ala Ser Thr Lys Leu Ser Pro
530 535 540
Trp Asp Lys Ser Ser His Ser Ala His Arg Val Asp Leu Ser Ala Val
545 550 555 560

<210> 48

<211> 394

<212> PRT

<213> Homo sapiens

<400> 48

Met	Thr	Ala	Gly	Arg	Ser	Gln	Glu	Arg	Arg	Ala	Gln	Glu	Met	Gly	Arg
1				5					10					15	
Gly	Ser	Val	Gln	Gly	Leu	Asp	Leu	Lys	Gly	Asp	Leu	Glu	Phe	Phe	Thr
			20					25					30		
Ala	Pro	Met	Leu	Ser	Leu	Arg	Ser	Phe	Val	Phe	Val	Gly	Val	Gly	Ser
		35					40						45		
Gly	Leu	Thr	Ser	Ser	His	Ile	Pro	Ala	Gln	Arg	Trp	Ala	Glu	Trp	Gly
	50					55					60				
Gln	Cys	Leu	Ala	Pro	Pro	Ala	Arg	Ser	Leu	Leu	Thr	Ser	Gly	Ser	Leu
65					70					75				80	
Cys	Cys	Pro	Arg	Thr	Met	Asn	Gly	Thr	Tyr	Asn	Thr	Cys	Gly	Ser	Ser
			85					90					95		
Asp	Leu	Thr	Trp	Pro	Pro	Ala	Ile	Lys	Leu	Gly	Phe	Tyr	Ala	Tyr	Leu
			100					105					110		
Gly	Val	Leu	Leu	Val	Leu	Gly	Leu	Leu	Leu	Asn	Ser	Leu	Ala	Leu	Trp
		115					120					125			
Val	Phe	Cys	Cys	Arg	Met	Gln	Gln	Trp	Thr	Glu	Thr	Arg	Ile	Tyr	Met
	130					135						140			
Thr	Asn	Leu	Ala	Val	Ala	Asp	Leu	Cys	Leu	Leu	Cys	Thr	Leu	Pro	Phe
145					150					155				160	
Val	Leu	His	Ser	Leu	Arg	Asp	Thr	Ser	Asp	Thr	Pro	Leu	Cys	Gln	Leu
				165						170				175	
Ser	Gln	Gly	Ile	Tyr	Leu	Thr	Asn	Arg	Tyr	Met	Ser	Ile	Ser	Leu	Val
		180						185					190		
Thr	Ala	Ile	Ala	Val	Asp	Arg	Tyr	Val	Ala	Val	Arg	His	Pro	Leu	Arg
		195						200					205		

Ala Arg Gly Leu Arg Ser Pro Arg Gln Ala Ala Ala Val Cys Ala Val
 210 215 220
 Leu Trp Val Leu Val Ile Gly Ser Leu Val Ala Arg Trp Leu Leu Gly
 225 230 235 240
 Ile Gln Glu Gly Gly Phe Cys Phe Arg Ser Thr Arg His Asn Phe Asn
 245 250 255
 Ser Met Ala Phe Pro Leu Leu Gly Phe Tyr Leu Pro Leu Ala Val Val
 260 265 270
 Val Phe Cys Ser Leu Lys Val Val Thr Ala Leu Ala Gln Arg Pro Pro
 275 280 285
 Thr Asp Val Gly Gln Ala Glu Ala Thr Arg Lys Ala Ala Arg Met Val
 290 295 300
 Trp Ala Asn Leu Leu Val Phe Val Val Cys Phe Leu Pro Leu His Val
 305 310 315 320
 Gly Leu Thr Val Arg Leu Ala Val Gly Trp Asn Ala Cys Ala Leu Leu
 325 330 335
 Glu Met Ile Arg Arg Ala Leu Tyr Ile Thr Ser Lys Leu Ser Asp Ala
 340 345 350
 Asn Cys Cys Leu Asp Ala Ile Cys Tyr Tyr Tyr Met Ala Lys Glu Phe
 355 360 365
 Gln Glu Ala Ser Ala Leu Ala Val Ala Pro Arg Ala Lys Ala His Lys
 370 375 380
 Ser Gln Asp Ser Leu Cys Val Thr Leu Ala
 385 390

<210> 49

<211> 422

<212> PRT

<213> Homo sapiens

<400> 49

Met Phe Arg Ala Gly Ala Glu Gly Ala Glu Lys Glu Pro Ser Pro Arg
 1 5 10 15
 Pro Glu Cys Arg Ala Asp Pro Gly Pro Gly Leu Gly Phe Pro Leu Gly
 20 25 30
 Ser Gly Leu Pro Trp Pro Ser Leu Leu Glu Ser Pro Gly Gly Arg Ile
 35 40 45
 Leu Asp Ile Pro Cys Lys Val Cys Gly Asp Arg Ser Ser Gly Lys His
 50 55 60
 Tyr Gly Val Tyr Ala Cys Asp Gly Cys Ser Gly Phe Phe Lys Arg Ser
 65 70 75 80
 Ile Arg Arg Asn Arg Thr Tyr Val Cys Lys Ser Gly Asn Gln Gly Gly
 85 90 95
 Cys Pro Val Asp Lys Thr His Arg Asn Gln Cys Arg Ala Cys Arg Leu
 100 105 110
 Lys Lys Cys Leu Glu Val Asn Met Asn Lys Asp Ala Val Gln His Glu
 115 120 125
 Arg Gly Pro Arg Thr Ser Thr Ile Arg Lys Gln Val Ala Leu Tyr Phe
 130 135 140
 Arg Gly His Lys Glu Glu Asn Gly Ala Ala Ala His Phe Pro Ser Ala
 145 150 155 160
 Ala Leu Pro Ala Pro Ala Phe Phe Thr Ala Val Thr Gln Leu Glu Pro
 165 170 175
 His Gly Leu Glu Leu Ala Ala Val Ser Thr Thr Pro Glu Arg Gln Thr
 180 185 190
 Leu Val Ser Leu Ala Gln Pro Thr Pro Lys Tyr Pro His Glu Val Asn
 195 200 205

Gly Thr Pro Met Tyr Leu Tyr Glu Val Ala Thr Glu Ser Val Cys Glu
210 215 220
Ser Ala Ala Arg Leu Leu Phe Met Ser Ile Lys Trp Ala Lys Ser Val
225 230 235 240
Pro Ala Phe Ser Thr Leu Ser Leu Gln Asp Gln Leu Met Leu Leu Glu
245 250 255
Asp Ala Trp Arg Glu Leu Phe Val Leu Gly Ile Ala Gln Trp Ala Ile
260 265 270
Pro Val Asp Ala Asn Thr Leu Leu Ala Val Ser Gly Met Asn Gly Asp
275 280 285
Asn Thr Asp Ser Gln Lys Leu Asn Lys Ile Ile Ser Glu Ile Gln Ala
290 295 300
Leu Gln Glu Val Val Ala Arg Phe Arg Gln Leu Arg Leu Asp Ala Thr
305 310 315 320
Glu Phe Ala Cys Leu Lys Cys Ile Val Thr Phe Lys Ala Val Pro Thr
325 330 335
His Ser Gly Ser Glu Leu Arg Ser Phe Arg Asn Ala Ala Ala Ile Ala
340 345 350
Ala Leu Gln Asp Glu Ala Gln Leu Thr Leu Asn Ser Tyr Ile His Thr
355 360 365
Arg Tyr Pro Thr Gln Pro Cys Arg Phe Gly Lys Leu Leu Leu Leu Leu
370 375 380
Pro Ala Leu Arg Ser Ile Ser Pro Ser Thr Ile Glu Glu Val Phe Phe
385 390 395 400
Lys Lys Thr Ile Gly Asn Val Pro Ile Thr Arg Leu Leu Ser Asp Met
405 410 415
Tyr Lys Ser Ser Asp Ile
420

<210> 50

<211> 102

<212> PRT

<213> Homo sapiens

<400> 50

Met Ser Asn Trp Pro Leu Phe Cys Phe Lys Lys Gln Ser Leu Lys Gln

1

5

10

15

Glu Glu Asp Arg Pro Lys Arg Arg Gly Pro Ala Tyr Val Met Glu Leu

20

25

30

Pro Lys Leu Lys Leu Ser Gly Val Val Arg Leu Ser Ser Tyr Ser Ser

35

40

45

Pro Thr Leu Gln Ser Val Leu Gly Ser Gly Thr Asn Gly Arg Val Pro

50

55

60

Val Leu Arg Pro Leu Lys Cys Ile Pro Ala Ser Lys Lys Thr Asp Pro

65

70

75

80

Gln Lys Asp Leu Lys Pro Ala Pro Gln Gln Cys Arg Leu Pro Thr Ile

85

90

95

Val Arg Lys Gly Gly Arg

100

<210> 51

<211> 205

<212> PRT

<213> Homo sapiens

<400> 51

Met Leu Pro Ser Gly Gln Pro Ile Gln Pro Leu Pro Arg Thr Cys Pro
 1 5 10 15
 Arg Arg Ser Gly Pro Ala Ala Ala Pro Arg Ala Arg Thr Pro Gly Thr
 20 25 30
 Gly Pro Gly Ala Ala Thr Pro His Ala Pro Gln Ala Thr Pro Pro Ser
 35 40 45
 Ala Val Ala Ala Arg Gly Pro Arg His Pro Pro Ser Glu Leu Ala Arg
 50 55 60
 Asp Pro Ala Asn Gly Pro Arg Gly Gly Gly Thr Leu Arg Asp Pro Thr
 65 70 75 80
 Ala Ala Val Gly Arg Ala Pro Gly Ala Pro Gly Gly Ser Arg Arg Gly
 85 90 95
 Arg His Ser Pro Ala Glu Lys Arg Arg Gly Arg Arg Ala Arg Ser Gly
 100 105 110
 Arg Ser Arg Ser Pro Trp Cys Pro Pro Ile Gln Glu Ala Pro Gly Cys
 115 120 125
 Pro Ala Leu Gly Pro Arg Pro Arg Ser Ala Arg Ser Pro Arg Ala Ala
 130 135 140
 Ala Ser Thr Pro Pro Pro Pro His Pro Val Pro Pro Ser Ser Pro Gln
 145 150 155 160
 Pro Pro Ala Gln Pro Gln Pro Arg Val Thr Ala Ala Leu Arg Leu Pro
 165 170 175
 Ser His Pro Arg Leu Arg Pro Leu Gly Ala Ser Arg Arg Pro Glu Arg
 180 185 190
 Gly Thr Leu Leu Lys Gly Ala Met Ala Arg Pro Leu Gly
 195 200 205

<211> 213

<212> PRT

<213> Homo sapiens

<400> 52

Met Glu Leu His Asp Arg Val Leu Asp Val Gln Leu Cys Ala Cys Ser

1 5 10 15

Leu Leu Leu His Leu Leu Gly Gln Gly Ile Ile Val Asn Lys Ala Pro

20 25 30

Leu Glu Lys Val Pro Asp Leu Ile Ser Gln Val Leu Ala Thr Tyr Pro

35 40 45

Ala Asp Gly Glu Met Ala Glu Ala Ser Cys Gly Val Phe Trp Leu Leu

50 55 60

Ser Leu Leu Gly Cys Ile Lys Glu Gln Gln Phe Glu Gln Val Val Ala

65 70 75 80

Leu Leu Leu Gln Ser Ile Arg Leu Cys Gln Asp Arg Ala Leu Leu Val

85 90 95

Asn Asn Ala Tyr Arg Gly Leu Ala Ser Leu Val Lys Val Ser Glu Leu

100 105 110

Ala Ala Phe Lys Val Val Val Gln Glu Glu Gly Gly Ser Gly Leu Ser

115 120 125

Leu Ile Lys Glu Thr Tyr Gln Leu His Arg Asp Asp Pro Glu Val Val

130 135 140

Glu Asn Val Gly Met Leu Leu Val His Leu Ala Ser Tyr Glu Glu Ile

145 150 155 160

Leu Pro Glu Leu Val Ser Ser Ser Met Lys Ala Leu Leu Gln Glu Ile

165 170 175

Lys Glu Arg Phe Thr Ser Ser Leu Val Ser Asp Ser Ser Ala Phe Ser

180 185 190

Lys Pro Gly Leu Pro Pro Gly Gly Ser Pro Gln Leu Gly Cys Thr Thr
 195 200 205
 Ser Gly Gly Leu Glu
 210

<210> 53

<211> 199

<212> PRT

<213> Homo sapiens

<400> 53

Met Gln Gly Leu Trp Phe Arg Ala Val Ile Cys Lys Gln Lys Phe Val
 1 5 10 15
 Ala Leu Val Ser Arg Arg Leu Gln Asp Thr His Met Ala Gln Leu Pro
 20 25 30
 Asp Lys Met Ala Ala Asn Gln Ile Cys Leu Pro Phe His Pro His Thr
 35 40 45
 Asp Val Cys Leu His Arg Asn Ser Phe Ser Met His Phe Ile Phe Ser
 50 55 60
 Asp Thr Ala Val Leu Leu Phe Asp Phe Trp Ser Val His Ser Pro Ala
 65 70 75 80
 Gly Met Ala Leu Ser Val Leu Val Leu Leu Leu Leu Ala Val Leu Tyr
 85 90 95
 Glu Gly Ile Lys Val Gly Lys Ala Lys Leu Leu Asn Gln Val Leu Val
 100 105 110
 Asn Leu Pro Thr Ser Ile Ser Gln Gln Thr Ile Ala Glu Thr Asp Gly
 115 120 125
 Asp Ser Ala Gly Ser Asp Ser Phe Pro Val Gly Arg Thr His His Arg

130 135 140
 Trp Tyr Leu Cys His Phe Gly Gln Ser Leu Ile His Val Ile Gln Val
 145 150 155 160
 Val Ile Gly Tyr Phe Ile Met Leu Ala Val Met Ser Tyr Asn Thr Trp
 165 170 175
 Ile Phe Leu Gly Val Val Leu Gly Ser Ala Val Gly Tyr Tyr Leu Ala
 180 185 190
 Tyr Pro Leu Leu Ser Thr Ala
 195

<210> 54

<211> 600

<212> PRT

<213> Homo sapiens

<400> 54

Met Ala Ala Ser Ala Gln Val Ser Val Thr Phe Glu Asp Val Ala Val
 1 5 10 15
 Thr Phe Thr Gln Glu Glu Trp Gly Gln Leu Asp Ala Ala Gln Arg Thr
 20 25 30
 Leu Tyr Gln Glu Val Met Leu Glu Thr Cys Gly Leu Leu Met Ser Leu
 35 40 45
 Gly Cys Pro Leu Phe Lys Pro Glu Leu Ile Tyr Gln Leu Asp His Arg
 50 55 60
 Gln Glu Leu Trp Met Ala Thr Lys Asp Leu Ser Gln Ser Ser Tyr Pro
 65 70 75 80
 Gly Asp Asn Thr Lys Pro Lys Thr Thr Glu Pro Thr Phe Ser His Leu
 85 90 95

Ala Leu Pro Glu Glu Val Leu Leu Gln Glu Gln Leu Thr Gln Gly Ala
 100 105 110
 Ser Lys Asn Ser Gln Leu Gly Gln Ser Lys Asp Gln Asp Gly Pro Ser
 115 120 125
 Glu Met Gln Glu Val His Leu Lys Ile Gly Ile Gly Pro Gln Arg Gly
 130 135 140
 Lys Leu Leu Glu Lys Met Ser Ser Glu Arg Asp Gly Leu Gly Ser Asp
 145 150 155 160
 Asp Gly Val Cys Thr Lys Ile Thr Gln Lys Gln Val Ser Thr Glu Gly
 165 170 175
 Asp Leu Tyr Glu Cys Asp Ser His Gly Pro Val Thr Asp Ala Leu Ile
 180 185 190
 Arg Glu Glu Lys Asn Ser Tyr Lys Cys Glu Glu Cys Gly Lys Val Phe
 195 200 205
 Lys Lys Asn Ala Leu Leu Val Gln His Glu Arg Ile His Thr Gln Val
 210 215 220
 Lys Pro Tyr Glu Cys Thr Glu Cys Gly Lys Thr Phe Ser Lys Ser Thr
 225 230 235 240
 His Leu Leu Gln His Leu Ile Ile His Thr Gly Glu Lys Pro Tyr Lys
 245 250 255
 Cys Met Glu Cys Gly Lys Ala Phe Asn Arg Arg Ser His Leu Thr Arg
 260 265 270
 His Gln Arg Ile His Ser Gly Glu Lys Pro Tyr Lys Cys Ser Glu Cys
 275 280 285
 Gly Lys Ala Phe Thr His Arg Ser Thr Phe Val Leu His His Arg Ser
 290 295 300
 His Thr Gly Glu Lys Pro Phe Val Cys Lys Glu Cys Gly Lys Ala Phe
 305 310 315 320
 Arg Asp Arg Pro Gly Phe Ile Arg His Tyr Ile Ile His Thr Gly Glu

325	330	335
Lys Pro Tyr Glu Cys Ile Glu Cys Gly	Lys Ala Phe Asn Arg Arg Ser	
340	345	350
Tyr Leu Thr Trp His Gln Gln Ile His Thr Gly Val Lys Pro Phe Glu		
355	360	365
Cys Asn Glu Cys Gly Lys Ala Phe Cys Glu Ser Ala Asp Leu Ile Gln		
370	375	380
His Tyr Ile Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Met Glu Cys		
385	390	395
Gly Lys Ala Phe Asn Arg Arg Ser His Leu Lys Gln His Gln Arg Ile		
405	410	415
His Thr Gly Glu Lys Pro Tyr Glu Cys Ser Glu Cys Gly Lys Ala Phe		
420	425	430
Thr His Cys Ser Thr Phe Val Leu His Lys Arg Thr His Thr Gly Glu		
435	440	445
Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys Ala Phe Ser Asp Arg Ala		
450	455	460
Asp Leu Ile Arg His Leu Ser Ile His Thr Gly Glu Lys Pro Tyr Glu		
465	470	475
Cys Val Glu Cys Gly Lys Ala Phe Asn Arg Ser Ser His Leu Thr Arg		
485	490	495
His Gln Gln Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Ile Gln Cys		
500	505	510
Gly Lys Ala Phe Cys Arg Ser Ala Asn Leu Ile Arg His Ser Ile Ile		
515	520	525
His Thr Gly Glu Lys Pro Tyr Glu Cys Ser Glu Cys Gly Lys Ala Phe		
530	535	540
Asn Arg Gly Ser Ser Leu Thr His His Gln Arg Ile His Thr Gly Arg		
545	550	555
		560

Asn Pro Thr Ile Val Thr Asp Val Gly Arg Pro Phe Met Thr Ala Gln
 565 570 575
 Thr Ser Val Asn Ile Gln Glu Leu Leu Leu Gly Lys Glu Phe Leu Asn
 580 585 590
 Ile Thr Thr Glu Glu Asn Leu Trp
 595 600

<210> 55

<211> 243

<212> PRT

<213> Homo sapiens

<400> 55

Met Lys Glu Ala Leu Ser Thr Glu Arg Gly Lys Thr Leu Val Gln Lys
 1 5 10 15
 Lys Pro Thr Met Tyr Pro Glu Trp Lys Ser Thr Phe Asp Ala His Ile
 20 25 30
 Tyr Glu Gly Arg Val Ile Gln Ile Val Leu Met Arg Ala Ala Glu Glu
 35 40 45
 Pro Val Ser Glu Val Thr Val Gly Val Ser Val Leu Ala Glu Arg Cys
 50 55 60
 Lys Lys Asn Asn Gly Lys Ala Glu Phe Trp Leu Asp Leu Gln Pro Gln
 65 70 75 80
 Ala Lys Val Leu Met Ser Val Gln Tyr Phe Leu Glu Asp Val Asp Cys
 85 90 95
 Lys Gln Ser Met Arg Ser Glu Asp Glu Ala Lys Phe Pro Thr Met Asn
 100 105 110
 Arg Arg Gly Ala Ile Lys Gln Ala Lys Ile His Tyr Ile Lys Asn His

<400> 56

Met Phe Phe Ser Phe Val Lys Lys Gly Ile Pro Leu Ile Phe Glu Arg
1 5 10 15
Glu Glu Ala Ile Lys Glu Gln Tyr Thr Val Arg Arg Phe Arg Gly Pro
20 25 30

Ser Cys Thr His Glu Leu Leu Arg Leu Pro Thr Ser Ser Arg Cys Arg
 35 40 45
 Pro Leu Arg Val Leu His Gly Glu Pro Ser Gln Pro Arg Trp Leu Pro
 50 55 60
 Asp Arg Glu Glu Pro Gln Val Gln Ala Leu Gln Gln Leu Gln Gly Ala
 65 70 75 80
 Ala Arg Val Phe Met Pro Leu Gln Ala Leu Asp Ser Ala Pro Lys Pro
 85 90 95
 Leu Lys Gly Gln Ala Gln Ala Pro Gln Arg Leu Gln Gly Ala Ala Arg
 100 105 110
 Val Phe Met Pro Leu Gln Ala Gln Val Lys Ala Lys Ala Ser Lys Pro
 115 120 125
 Leu Gln Met Gln Ile Lys Ala Pro Pro Arg Leu Arg Arg Ala Ala Arg
 130 135 140
 Val Leu Met Pro Leu Gln Ala Gln Val Arg Ala Pro Arg Leu Leu Gln
 145 150 155 160
 Val Gln Ser Gln Val Ser Lys Lys Gln Gln Ala Gln Thr Gln Thr Ser
 165 170 175
 Glu Pro Gln Asp Leu Asp Gln Val Pro Glu Glu Phe Gln Gly Gln Asp
 180 185 190
 Gln Val Pro Glu Gln Gln Arg Gln Gly Gln Ala Pro Glu Gln Gln Gln
 195 200 205
 Arg His Asn Gln Val Pro Glu Gln Glu Leu Glu Gln Asn Gln Ala Pro
 210 215 220
 Glu Gln Pro Glu Val Gln Glu Gln Ala Ala Glu Pro Ala Gln Ala Glu
 225 230 235 240
 Thr Glu Ala Glu Glu Pro Glu Ser Leu Arg Val Asn Ala Gln Val Phe
 245 250 255
 Leu Pro Leu Leu Ser Gln Asp His His Val Leu Leu Pro Leu His Leu

260 265 270
Asp Thr Gln Val Leu Ile Pro Val Glu Gly Gln Thr Glu Gly Ser Pro
275 280 285
Gln Ala Gln Ala Trp Thr Leu Glu Pro Pro Gln Ala Ile Gly Ser Val
290 295 300
Gln Ala Leu Ile Glu Gly Leu Ser Arg Asp Leu Leu Arg Ala Pro Asn
305 310 315 320
Ser Asn Asn Ser Lys Pro Leu Gly Pro Leu Gln Thr Leu Met Glu Asn
325 330 335
Leu Ser Ser Asn Arg Phe Tyr Ser Gln Pro Glu Gln Ala Arg Glu Lys
340 345 350
Lys Ser Lys Val Ser Thr Leu Arg Gln Ala Leu Ala Lys Arg Leu Ser
355 360 365
Pro Lys Arg Phe Arg Ala Lys Ser Ser Trp Arg Pro Glu Lys Leu Glu
370 375 380
Leu Ser Asp Leu Glu Ala Arg Arg Gln Arg Arg Gln Arg Arg Trp Glu
385 390 395 400
Asp Ile Phe Asn Gln His Glu Glu Glu Leu Arg Gln Val Asp Lys Asp
405 410 415
Lys Glu Asp Glu Ser Ser Asp Asn Asp Glu Val Phe His Ser Ile Gln
420 425 430
Ala Glu Val Gln Ile Glu Pro Leu Lys Pro Tyr Ile Ser Asn Pro Lys
435 440 445
Lys Ile Glu Val Gln Glu Arg Ser Pro Ser Val Pro Asn Asn Gln Asp
450 455 460
His Ala His His Val Lys Phe Ser Ser Ser Val Pro Gln Arg Ser Leu
465 470 475 480
Leu Glu Gln Ala Gln Lys Pro Ile Asp Ile Arg Gln Arg Ser Ser Gln
485 490 495

Asn Arg Gln Asn Trp Leu Ala Ala Ser Glu Ser Ser Ser Glu Glu Glu
 500 505 510
 Ser Pro Val Thr Gly Arg Arg Ser Gln Ser Ser Pro Pro Tyr Ser Thr
 515 520 525
 Ile Asp Gln Lys Leu Leu Val Asp Ile His Val Pro Asp Gly Phe Lys
 530 535 540
 Val Gly Lys Ile Ser Pro Pro Val Tyr Leu Thr Asn Glu Trp Val Gly
 545 550 555 560
 Tyr Asn Ala Leu Ser Glu Ile Phe Arg Asn Asp Trp Leu Thr Pro Ala
 565 570 575
 Pro Val Ile Gln Pro Pro Glu Glu Asp Gly Asp Tyr Val Glu Leu Tyr
 580 585 590
 Asp Ala Ser Ala Asp Thr Asp Gly Asp Asp Asp Asp Glu Ser Asn Asp
 595 600 605
 Thr Phe Glu Asp Thr Tyr Asp His Ala Asn Gly Asn Asp Asp Leu Asp
 610 615 620
 Asn Gln Val Asp Gln Ala Asn Asp Val Cys Lys Asp His Asp Asp Asp
 625 630 635 640
 Asn Asn Lys Phe Val Asp Asp Val Asn Asn Asn Tyr Tyr Glu Ala Pro
 645 650 655
 Ser Cys Pro Arg Ala Ser Tyr Gly Arg Asp Gly Ser Cys Lys Gln Asp
 660 665 670
 Gly Tyr Asp Gly Ser Arg Gly Lys Glu Glu Ala Tyr Arg Gly Tyr Gly
 675 680 685
 Ser His Thr Ala Asn Arg Ser His Gly Gly Ser Ala Ala Ser Glu Asp
 690 695 700
 Asn Ala Ala Ile Gly Asp Gln Glu Glu His Ala Ala Asn Ile Gly Ser
 705 710 715 720
 Glu Arg Arg Gly Ser Glu Gly Asp Gly Gly Lys Gly Val Val Arg Thr

725	730	735
Ser Glu Glu Ser Gly Ala Leu Gly Leu Asn Gly Glu Glu Asn Cys Ser		
740	745	750
Glu Thr Asp Gly Pro Gly Leu Lys Arg Pro Ala Ser Gln Asp Phe Glu		
755	760	765
Tyr Leu Gln Glu Glu Pro Gly Gly Gly Asn Glu Ala Ser Asn Ala Ile		
770	775	780
Asp Ser Gly Ala Ala Pro Ser Ala Pro Asp His Glu Ser Asp Asn Lys		
785	790	795
800		
Asp Ile Ser Glu Ser Ser Thr Gln Ser Asp Phe Ser Ala Asn His Ser		
805	810	815
Ser Pro Ser Lys Gly Ser Gly Met Ser Ala Asp Ala Asn Phe Ala Ser		
820	825	830
Ala Ile Leu Tyr Ala Gly Phe Val Glu Val Pro Glu Glu Ser Pro Lys		
835	840	845
Gln Pro Ser Glu Val Asn Val Asn Pro Leu Tyr Val Ser Pro Ala Cys		
850	855	860
Lys Lys Pro Leu Ile His Met Tyr Glu Lys Glu Phe Thr Ser Glu Ile		
865	870	875
880		
Cys Cys Gly Ser Leu Trp Gly Val Asn Leu Leu Leu Gly Thr Arg Ser		
885	890	895
Asn Leu Tyr Leu Met Asp Arg Ser Gly Lys Ala Asp Ile Thr Lys Leu		
900	905	910
Ile Arg Arg Arg Pro Phe Arg Gln Ile Gln Val Leu Glu Pro Leu Asn		
915	920	925
Leu Leu Ile Thr Ile Ser Gly His Lys Asn Arg Leu Arg Val Tyr His		
930	935	940
Leu Thr Trp Leu Arg Asn Lys Ile Leu Asn Asn Asp Pro Glu Ser		
945	950	955

<210> 57

<211> 523

<212> PRT

<213> Homo sapiens

<400> 57

Met Thr Pro Gln Ser Leu Leu Gln Thr Thr Leu Phe Leu Leu Ser Leu

1 5 10 15

Leu Phe Leu Val Gln Gly Ala His Gly Arg Gly His Arg Glu Asp Phe

20 25 30

Arg Phe Cys Ser Gln Gln Leu Gln Ser Leu Glu Ser Lys Leu Thr Ser

35 40 45

Val Arg Phe Met Gly Asp Met Val Ser Phe Glu Glu Asp Arg Ile Asn

50 55 60

Ala Thr Val Trp Lys Leu Gln Pro Thr Ala Gly Leu Gln Asp Leu His

65 70 75 80

Ile His Ser Arg Gln Glu Glu Glu Gln Ser Glu Ile Met Glu Tyr Ser

85 90 95

Val Leu Leu Pro Arg Thr Leu Phe Gln Arg Thr Lys Gly Arg Arg Gly

100 105 110

Glu Ala Glu Lys Arg Leu Leu Leu Val Asp Phe Ser Ser Gln Ala Leu

115 120 125

Phe Gln Asp Lys Asn Ser Ser His Val Leu Gly Glu Lys Val Leu Gly

130 135 140

Ile Val Val Gln Asn Thr Lys Val Ala Asn Leu Thr Glu Pro Val Val

145 150 155 160

Leu Thr Phe Gln His Gln Leu Gln Pro Lys Asn Val Thr Leu Gln Cys

	165		170		175
Val Phe Trp	Val Glu Asp Pro Thr Leu Ser Ser Pro Gly His Trp Ser				
	180		185		190
Ser Ala Gly Cys Glu Thr Val Arg Arg Glu Thr Gln Thr Ser Cys Phe					
	195		200		205
Cys Asn His Leu Thr Tyr Phe Ala Val Leu Met Val Ser Ser Val Glu					
	210		215		220
Val Asp Ala Val His Lys His Tyr Leu Ser Leu Leu Ser Tyr Val Gly					
225		230		235	240
Cys Val Val Ser Ala Leu Ala Cys Leu Val Thr Ile Ala Ala Tyr Leu					
	245		250		255
Cys Ser Arg Val Pro Leu Pro Cys Arg Arg Lys Pro Arg Asp Tyr Thr					
	260		265		270
Ile Lys Val His Met Asn Leu Leu Leu Ala Val Phe Leu Leu Asp Thr					
	275		280		285
Ser Phe Leu Leu Ser Glu Pro Val Ala Leu Thr Gly Ser Glu Ala Gly					
	290		295		300
Cys Arg Ala Ser Ala Ile Phe Leu His Phe Ser Leu Leu Thr Cys Leu					
305		310		315	320
Ser Trp Met Gly Leu Glu Gly Tyr Asn Leu Tyr Arg Leu Val Val Glu					
	325		330		335
Val Phe Gly Thr Tyr Val Pro Gly Tyr Leu Leu Lys Leu Ser Ala Met					
	340		345		350
Gly Trp Gly Phe Pro Ile Phe Leu Val Thr Leu Val Ala Leu Val Asp					
	355		360		365
Val Asp Asn Tyr Gly Pro Ile Ile Leu Ala Val His Arg Thr Pro Glu					
	370		375		380
Gly Val Ile Tyr Pro Ser Arg Cys Trp Ile Arg Asp Ser Leu Val Ser					
385		390		395	400

Tyr Ile Thr Asn Leu Gly Leu Phe Ser Leu Val Phe Leu Phe Asn Met
 405 410 415
 Ala Met Leu Ala Thr Met Val Val Gln Ile Leu Arg Leu Arg Pro His
 420 425 430
 Thr Gln Lys Trp Ser His Val Leu Thr Leu Leu Gly Leu Ser Leu Val
 435 440 445
 Leu Gly Leu Pro Trp Ala Leu Ile Phe Phe Ser Phe Ala Ser Gly Thr
 450 455 460
 Phe Gln Leu Val Val Leu Tyr Leu Phe Ser Ile Ile Thr Ser Phe Gln
 465 470 475 480
 Gly Phe Leu Ile Phe Ile Trp Tyr Trp Ser Met Arg Leu Gln Ala Arg
 485 490 495
 Gly Gly Pro Ser Pro Leu Lys Ser Asn Ser Asp Ser Ala Arg Leu Pro
 500 505 510
 Ile Ser Ser Gly Ser Thr Ser Ser Ser Arg Ile
 515 520

<210> 58

<211> 576

<212> PRT

<213> Homo sapiens

<400> 58

Met Val Ser His Phe Met Gly Ser Leu Ser Val Leu Cys Phe Leu Leu
 1 5 10 15
 Leu Leu Gly Phe Gln Phe Val Cys Pro Gln Pro Ser Thr Gln His Arg
 20 25 30
 Lys Val Pro Gln Arg Met Ala Ala Glu Gly Ala Pro Glu Asp Asp Gly

35	40	45
Gly Gly Gly Ala Pro Gly Val Trp Gly Ala Trp Gly Pro Trp Ser Ala		
50	55	60
Cys Ser Arg Ser Cys Ser Gly Gly Val Met Glu Gln Thr Arg Pro Cys		
65	70	75
80		
Leu Pro Arg Ser Tyr Arg Leu Arg Gly Gly Gln Arg Pro Gly Ala Pro		
85	90	95
Ala Arg Ala Phe Ala Asp His Ala Val Ser Ala Val Arg Thr Ser Val		
100	105	110
Pro Leu His Arg Ser Arg Asp Glu Thr Pro Ala Leu Ala Gly Thr Asp		
115	120	125
Ala Ser Arg Gln Gly Pro Thr Val Leu Arg Gly Ser Arg His Pro Gln		
130	135	140
Pro Gln Gly Leu Glu Val Thr Gly Asp Arg Arg Ser Arg Thr Arg Gly		
145	150	155
160		
Thr Ile Gly Pro Gly Lys Tyr Gly Tyr Gly Lys Ala Pro Tyr Ile Leu		
165	170	175
Pro Leu Gln Thr Asp Thr Ala His Thr Pro Gln Arg Leu Arg Arg Gln		
180	185	190
Lys Leu Ser Ser Arg His Ser Arg Ser Gln Gly Ala Ser Ser Ala Arg		
195	200	205
His Gly Tyr Ser Ser Pro Ala His Gln Val Pro Gln His Gly Pro Leu		
210	215	220
Tyr Gln Ser Asp Ser Gly Pro Arg Ser Gly Leu Gln Ala Ala Glu Ala		
225	230	235
240		
Pro Ile Tyr Gln Leu Pro Leu Thr His Asp Gln Gly Tyr Pro Ala Ala		
245	250	255
Ser Ser Leu Phe His Ser Pro Glu Thr Ser Asn Asn His Gly Val Gly		
260	265	270

Thr	His	Gly	Ala	Thr	Gln	Ser	Phe	Ser	Gln	Pro	Ala	Arg	Ser	Thr	Ala				
																275	280	285	
Ile	Ser	Cys	Ile	Gly	Ala	Tyr	Arg	Gln	Tyr	Lys	Leu	Cys	Asn	Thr	Asn				
																290	295	300	
Val	Cys	Pro	Glu	Ser	Ser	Arg	Ser	Ile	Arg	Glu	Val	Gln	Cys	Ala	Ser				
																305	310	315	320
Tyr	Asn	Asn	Lys	Pro	Phe	Met	Gly	Arg	Phe	Tyr	Glu	Trp	Glu	Pro	Phe				
																325	330	335	
Ala	Glu	Val	Lys	Gly	Asn	Arg	Lys	Cys	Glu	Leu	Asn	Cys	Gln	Ala	Met				
																340	345	350	
Gly	Tyr	Arg	Phe	Tyr	Val	Arg	Gln	Ala	Glu	Lys	Val	Ile	Asp	Gly	Thr				
																355	360	365	
Pro	Cys	Asp	Gln	Asn	Gly	Thr	Ala	Ile	Cys	Val	Ser	Gly	Gln	Cys	Lys				
																370	375	380	
Ser	Ile	Gly	Cys	Asp	Asp	Tyr	Leu	Gly	Ser	Asp	Lys	Val	Val	Asp	Lys				
																385	390	395	400
Cys	Gly	Val	Cys	Gly	Gly	Asp	Asn	Thr	Gly	Cys	Gln	Val	Val	Ser	Gly				
																405	410	415	
Val	Phe	Lys	His	Ala	Leu	Thr	Ser	Leu	Gly	Tyr	His	Arg	Val	Val	Glu				
																420	425	430	
Ile	Ser	Glu	Gly	Ala	Thr	Lys	Ile	Asn	Ile	Thr	Glu	Met	Tyr	Lys	Ser				
																435	440	445	
Asn	Asn	Tyr	Leu	Ala	Leu	Arg	Ser	Arg	Ser	Gly	Arg	Ser	Ile	Ile	Asn				
																450	455	460	
Gly	Asn	Trp	Ala	Ile	Asp	Arg	Pro	Gly	Lys	Tyr	Glu	Gly	Gly	Gly	Thr				
																465	470	475	480
Met	Phe	Thr	Tyr	Lys	Arg	Pro	Asn	Glu	Ile	Ser	Ser	Thr	Ala	Gly	Glu				
																485	490	495	
Ser	Phe	Leu	Ala	Glu	Gly	Pro	Thr	Asn	Glu	Ile	Leu	Asp	Val	Tyr	Met				

500 505 510
Ile His Gln Gln Pro Asn Pro Gly Val His Tyr Glu Tyr Val Ile Met
515 520 525
Gly Thr Asn Ala Ile Ser Pro Gln Val Pro Pro His Arg Arg Pro Gly
530 535 540
Glu Pro Phe Asn Gly Gln Met Val Thr Glu Gly Arg Ser Gln Glu Glu
545 550 555 560
Gly Glu Gln Lys Gly Arg Asn Glu Glu Lys Glu Asp Leu Arg Gly Glu
565 570 575

<210> 59

<211> 682

<212> PRT

<213> Homo sapiens

<400> 59

Met Gly Val Ala Gly Arg Asn Arg Pro Gly Ala Ala Trp Ala Val Leu
1 5 10 15
Leu Leu Leu Leu Leu Leu Pro Pro Leu Leu Leu Leu Ala Gly Ala Val
20 25 30
Pro Pro Gly Arg Gly Arg Ala Ala Gly Pro Gln Glu Asp Val Asp Glu
35 40 45
Cys Ala Gln Gly Leu Asp Asp Cys His Ala Asp Ala Leu Cys Gln Asn
50 55 60
Thr Pro Thr Ser Tyr Lys Cys Ser Cys Lys Pro Gly Tyr Gln Gly Glu
65 70 75 80
Gly Arg Gln Cys Glu Asp Ile Asp Glu Cys Gly Asn Glu Leu Asn Gly
85 90 95

Gly Cys Val His Asp Cys Leu Asn Ile Pro Gly Asn Tyr Arg Cys Thr
 100 105 110
 Cys Phe Asp Gly Phe Met Leu Ala His Asp Gly His Asn Cys Leu Asp
 115 120 125
 Val Asp Glu Cys Leu Glu Asn Asn Gly Gly Cys Gln His Thr Cys Val
 130 135 140
 Asn Val Met Gly Ser Tyr Glu Cys Cys Cys Lys Glu Gly Phe Phe Leu
 145 150 155 160
 Ser Asp Asn Gln His Thr Cys Ile His Arg Ser Glu Glu Gly Leu Ser
 165 170 175
 Cys Met Asn Lys Asp His Gly Cys Ser His Ile Cys Lys Glu Ala Pro
 180 185 190
 Arg Gly Ser Val Ala Cys Glu Cys Arg Pro Gly Phe Glu Leu Ala Lys
 195 200 205
 Asn Gln Arg Asp Cys Ile Leu Thr Cys Asn His Gly Asn Gly Gly Cys
 210 215 220
 Gln His Ser Cys Asp Asp Thr Ala Asp Gly Pro Glu Cys Ser Cys His
 225 230 235 240
 Pro Gln Tyr Lys Met His Thr Asp Gly Arg Ser Cys Leu Glu Arg Glu
 245 250 255
 Asp Thr Val Leu Glu Val Thr Glu Ser Asn Thr Thr Ser Val Val Asp
 260 265 270
 Gly Asp Lys Arg Val Lys Arg Gln Leu Leu Met Glu Thr Cys Ala Val
 275 280 285
 Asn Asn Gly Gly Cys Asp Arg Thr Cys Lys Asp Thr Ser Thr Gly Val
 290 295 300
 His Cys Ser Cys Pro Val Gly Phe Thr Leu Gln Leu Asp Gly Lys Thr
 305 310 315 320
 Cys Lys Asp Ile Asp Glu Cys Gln Thr Arg Asn Gly Gly Cys Asp His

325	330	335
Phe Cys Lys Asn Ile Val Gly Ser Phe Asp Cys Gly Cys Lys Lys Gly		
340	345	350
Phe Lys Leu Leu Thr Asp Glu Lys Ser Cys Gln Asp Val Asp Glu Cys		
355	360	365
Ser Leu Asp Arg Thr Cys Asp His Ser Cys Ile Asn His Pro Gly Thr		
370	375	380
Phe Ala Cys Ala Cys Asn Arg Gly Tyr Thr Leu Tyr Gly Phe Thr His		
385	390	395
Cys Gly Asp Thr Asn Glu Cys Ser Ile Asn Asn Gly Gly Cys Gln Gln		
405	410	415
Val Cys Val Asn Thr Val Gly Ser Tyr Glu Cys Gln Cys His Pro Gly		
420	425	430
Tyr Lys Leu His Trp Asn Lys Lys Asp Cys Val Ala Ser Cys Asp Leu		
435	440	445
Ser Cys Ile Val Lys Arg Thr Glu Lys Arg Leu Arg Lys Ala Ile Arg		
450	455	460
Thr Leu Arg Lys Ala Val His Arg Glu Gln Phe His Leu Gln Leu Ser		
465	470	475
Gly Met Asn Leu Asp Val Ala Lys Lys Pro Pro Arg Thr Ser Glu Arg		
485	490	495
Gln Ala Glu Ser Cys Gly Val Gly Gln Gly His Ala Glu Asn Gln Cys		
500	505	510
Val Ser Cys Arg Ala Gly Thr Tyr Tyr Asp Gly Ala Arg Glu Arg Cys		
515	520	525
Ile Leu Cys Pro Asn Gly Thr Phe Gln Asn Glu Glu Gly Gln Met Thr		
530	535	540
Cys Glu Pro Cys Pro Arg Pro Gly Asn Ser Gly Ala Leu Lys Thr Pro		
545	550	555
		560

Glu Ala Trp Asn Met Ser Glu Cys Gly Gly Leu Cys Gln Pro Gly Glu
 565 570 575
 Tyr Ser Ala Asp Gly Phe Ala Pro Cys Gln Leu Cys Ala Leu Gly Thr
 580 585 590
 Phe Gln Pro Glu Ala Gly Arg Thr Ser Cys Phe Pro Cys Gly Gly Gly
 595 600 605
 Leu Ala Thr Lys His Gln Gly Ala Thr Ser Phe Gln Asp Cys Glu Thr
 610 615 620
 Arg Val Gln Cys Ser Pro Gly His Phe Tyr Asn Thr Thr Thr His Arg
 625 630 635 640
 Cys Ile Arg Cys Pro Val Gly Thr Tyr Gln Pro Glu Phe Gly Lys Asn
 645 650 655
 Asn Cys Val Ser Cys Pro Gly Asn Thr Thr Thr Asp Phe Asp Gly Ser
 660 665 670
 Thr Asn Ile Thr Gln Phe Cys Asp Asn Ile
 675 680

<210> 60

<211> 369

<212> PRT

<213> Homo sapiens

<400> 60

Met Ala Ser Gly Ala Gln Leu Pro Pro Gln Pro Ser Ser Ser Glu Val
 1 5 10 15
 Ser Ala Val Gln Ser Pro Gly Gly Arg Pro Gly Ala Gly Leu Glu Glu
 20 25 30
 Thr Ala Leu Gly Val Pro Leu Pro Pro Ser Pro Gly Glu Ala Pro Leu

35	40	45
Pro Arg Ser Asn Arg Ser Arg Cys Pro Gly Thr Arg Gln Pro Gly Ala		
50	55	60
Ala Ser Leu His Ala Ala Ser Ala Ala Val Pro Val Arg Pro Arg Arg		
65	70	75
Gly Thr Ala Pro Ala Gly Lys Thr Ala Asp Ala Val Pro Ala Ala Ala		
85	90	95
Pro Glu Gln Ala Pro Arg Pro Ala Pro Gln Ser Arg Lys Pro Arg Asn		
100	105	110
Leu Glu Gly Asp Leu Asp Glu Arg Arg Leu Leu Cys His Leu Gln Leu		
115	120	125
Ala Gln Asp Arg Glu Ala Arg Leu Trp Arg Gly Gly Lys Pro Gln Asp		
130	135	140
Glu Ile Cys Asp Ala Phe Glu Glu Val Val Leu Trp Leu Leu Arg Leu		
145	150	155
Gln Asn Thr Phe Tyr Phe Ser Gln Ser Thr Phe Asn Leu Ala Leu Thr		
165	170	175
Ile Phe Gly Arg Leu Leu Ile Ser Val Lys Val Lys Glu Lys Tyr Leu		
180	185	190
His Cys Ala Thr Ile Thr Ser Leu Arg Leu Ala Ala Lys Val Asn Glu		
195	200	205
Glu Glu Glu Phe Ile Pro Gln Val Lys Asp Phe Thr Lys His Tyr Gly		
210	215	220
Ser Asp Tyr Ser Pro Asn Glu Leu Leu Arg Met Glu Leu Ala Ile Leu		
225	230	235
Asp Arg Leu His Trp Asp Leu Tyr Ile Gly Thr Pro Leu Asp Phe Leu		
245	250	255
Thr Ile Phe His Ala Leu Val Val Leu Ser Trp Pro His Val Leu Glu		
260	265	270

Leu Leu Pro Gln Arg Asn Pro Ser Leu His Val Ala Ser Leu Thr Arg
 275 280 285
 Gln Leu Gln His Cys Met Ala Gly His Gln Leu Leu Gln Phe Lys Gly
 290 295 300
 Ser Thr Leu Ala Leu Val Ile Ile Thr Leu Glu Leu Glu Arg Leu Met
 305 310 315 320
 Pro Gly Trp Cys Ala Pro Ile Ser Asp Leu Leu Lys Lys Ala Gln Val
 325 330 335
 Gly Asp Met Gln Tyr Ser Cys Cys Lys Glu Leu Val Met Gln Gln Leu
 340 345 350
 Arg Ser Leu Gln Ser Ser Ser Cys Thr Asp Asn Phe Val Ser Pro Ala
 355 360 365
 Asn

<210> 61

<211> 776

<212> PRT

<213> Homo sapiens

<400> 61

Met Ser Gly Lys Arg Gln Phe Phe Leu Lys Pro Leu Ala Leu Leu Thr
 1 5 10 15
 Cys Trp Pro Ser Thr Val Asn Gly Lys Asp Leu Ser Lys Ala Thr His
 20 25 30
 Glu Glu Ala Val Glu Ala Phe Arg Asn Ala Lys Glu Pro Ile Val Val
 35 40 45
 Gln Val Leu Arg Arg Thr Pro Leu Ser Arg Pro Ala Tyr Gly Met Ala

出証特 2 0 0 4 - 3 0 5 9 6 6 3

His Glu Lys Asp Ser Gly Val Gly Arg Thr Asp Glu Ser Leu Arg Asn
290 295 300
Asp Glu Ser Ser Glu Gln Glu Asn Ala Ala Glu Asp Pro Asn Ser Thr
305 310 315 320
Ser Leu Lys Ser Lys Arg Asp Leu Gly Gln Ser Gln Asp Thr Leu Gly
325 330 335
Ser Val Glu Leu Gln Tyr Asn Glu Ser Leu Val Ser Gly Glu Tyr Ile
340 345 350
Asp Ser Asp Cys Ile Gly Asn Pro Asp Glu Asp Cys Glu Arg Phe Arg
355 360 365
Gln Leu Leu Glu Leu Lys Cys Lys Ile Arg Asn His Gly Glu Tyr Asp
370 375 380
Leu Tyr Tyr Ser Ser Gly Thr Ile Glu Cys Asn Gln Gly Glu Gln Glu
385 390 395 400
Gly Val Glu His Glu Leu Gln Leu Leu Asn Glu Glu Leu Arg Asn Ile
405 410 415
Glu Leu Glu Cys Gln Asn Ile Met Gln Ala His Arg Leu Gln Lys Val
420 425 430
Thr Asp Gln Tyr Gly Asp Ile Trp Thr Leu His Asp Gly Gly Phe Arg
435 440 445
Asn Tyr Asn Thr Ser Ile Asp Met Gln Arg Gly Lys Leu Asp Asp Ile
450 455 460
Met Glu His Pro Glu Lys Ser Asp Lys Asp Ser Ser Ser Ala Tyr Asn
465 470 475 480
Thr Ala Glu Ser Cys Arg Ser Thr Pro Leu Thr Val Asp Arg Ser Pro
485 490 495
Asp Ser Ser Leu Pro Arg Val Ile Asn Leu Thr Asn Lys Lys Asn Leu
500 505 510
Arg Ser Thr Met Ala Ala Thr Gln Ser Ser Ser Gly Gln Ser Ser Lys

515 520 525
Glu Ser Thr Ser Thr Lys Ala Lys Thr Thr Glu Gln Gly Cys Ser Ala
530 535 540
Glu Ser Lys Glu Lys Val Leu Glu Gly Ser Lys Leu Pro Asp Gln Glu
545 550 555 560
Lys Ala Val Ser Glu His Ile Pro Tyr Leu Ser Pro Tyr His Ser Ser
565 570 575
Ser Tyr Arg Tyr Ala Asn Ile Pro Ala His Ala Arg His Tyr Gln Ser
580 585 590
Tyr Met Gln Leu Ile Gln Gln Lys Ser Ala Val Glu Tyr Ala Gln Ser
595 600 605
Gln Leu Ser Leu Val Ser Met Cys Lys Glu Ser Gln Lys Cys Ser Glu
610 615 620
Pro Lys Met Glu Trp Lys Val Lys Ile Arg Ser Asp Gly Thr Arg Tyr
625 630 635 640
Ile Thr Lys Arg Pro Val Arg Asp Arg Ile Leu Lys Glu Arg Ala Leu
645 650 655
Lys Ile Lys Glu Glu Arg Ser Gly Met Thr Thr Asp Asp Asp Thr Met
660 665 670
Ser Glu Met Lys Met Gly Arg Tyr Trp Ser Lys Glu Glu Arg Lys Gln
675 680 685
His Leu Val Arg Ala Lys Glu Gln Arg Arg Arg Arg Glu Phe Met Met
690 695 700
Arg Ser Arg Leu Glu Cys Leu Lys Glu Ser Pro Gln Ser Gly Ser Glu
705 710 715 720
Gly Lys Lys Glu Ile Asn Ile Ile Glu Leu Ser His Lys Lys Met Met
725 730 735
Lys Lys Arg Asn Lys Lys Ile Leu Asp Asn Trp Met Thr Ile Gln Glu
740 745 750

Leu Met Thr His Gly Ala Lys Ser Pro Asp Gly Thr Arg Val His Asn

755

760

765

Ala Phe Leu Ser Val Thr Thr Val

770

775

<210> 62

<211> 328

<212> PRT

<213> Homo sapiens

<400> 62

Met Lys Leu Trp Met Glu Ser His Leu Ile Val Pro Glu Thr Arg Pro

1

5

10

15

Ser Pro Arg Met Met Ser Asn Gln Thr Leu Val Thr Glu Phe Ile Leu

20

25

30

Gln Gly Phe Ser Glu His Pro Glu Tyr Arg Val Phe Leu Phe Ser Cys

35

40

45

Phe Leu Phe Leu Tyr Ser Gly Ala Leu Thr Gly Asn Val Leu Ile Thr

50

55

60

Leu Ala Ile Thr Phe Asn Pro Gly Leu His Ala Pro Met Tyr Phe Phe

65

70

75

80

Leu Leu Asn Leu Ala Thr Met Asp Ile Ile Cys Thr Ser Ser Ile Met

85

90

95

Pro Lys Ala Leu Ala Ser Leu Val Ser Glu Glu Ser Ser Ile Ser Tyr

100

105

110

Gly Gly Cys Met Ala Gln Leu Tyr Phe Leu Thr Trp Ala Ala Ser Ser

115

120

125

Glu Leu Leu Leu Leu Thr Val Met Ala Tyr Asp Arg Tyr Ala Ala Ile

130	135	140
Cys His Pro Leu His Tyr Ser Ser Met Met Ser Lys Val Phe Cys Ser		
145	150	155
Gly Leu Ala Thr Ala Val Trp Leu Leu Cys Ala Val Asn Thr Ala Ile		
165	170	175
His Thr Gly Leu Met Leu Arg Leu Asp Phe Cys Gly Pro Asn Val Ile		
180	185	190
Ile His Phe Phe Cys Gly Val Pro Pro Leu Leu Leu Leu Ser Cys Ser		
195	200	205
Ser Thr Tyr Val Asn Gly Val Met Ile Val Leu Ala Asp Ala Phe Tyr		
210	215	220
Gly Ile Val Asn Phe Leu Met Thr Ile Ala Ser Tyr Gly Phe Ile Val		
225	230	235
Ser Ser Ile Leu Lys Val Lys Thr Ala Trp Gly Arg Gln Lys Ala Phe		
245	250	255
Ser Thr Cys Ser Ser His Leu Thr Val Val Cys Met Tyr Tyr Thr Ala		
260	265	270
Val Phe Tyr Ala Tyr Ile Ser Pro Val Ser Gly Tyr Ser Ala Gly Lys		
275	280	285
Ser Lys Leu Ala Gly Leu Leu Tyr Thr Val Leu Ser Pro Thr Leu Asn		
290	295	300
Pro Leu Ile Tyr Thr Leu Arg Asn Lys Glu Val Lys Ala Ala Leu Arg		
305	310	315
Lys Leu Phe Pro Phe Phe Arg Asn		
325		

<210> 63

<211> 140

<212> PRT

<213> Homo sapiens

<400> 63

Met Val Arg Tyr Arg Val Arg Ser Leu Ser Glu Arg Ser His Glu Val

1 5 10 15

Tyr Arg Gln Gln Leu His Gly Gln Glu Gln Gly His His Gly Gln Glu

20 25 30

Glu Gln Gly Leu Ser Pro Glu His Val Glu Val Tyr Glu Arg Thr His

35 40 45

Gly Gln Ser His Tyr Arg Arg Arg His Cys Ser Arg Arg Arg Leu His

50 55 60

Arg Ile His Arg Arg Gln His Arg Ser Cys Arg Arg Arg Lys Arg Arg

65 70 75 80

Ser Cys Arg His Arg Arg Arg His Arg Arg Glu Ser Leu Gly Asp Pro

85 90 95

Leu Asn Gln Asn Phe Leu Ser Gln Lys Ala Ala Glu Pro Gly Arg Glu

100 105 110

His Ala Glu Gly Thr Lys Leu Pro Gly Pro Leu Thr Pro Ser Trp Lys

115 120 125

Leu Arg Lys Ser Arg Pro Lys His Gln Val Arg Pro

130 135 140

<210> 64

<211> 696

<212> PRT

<213> Homo sapiens

<400> 64

Met	Ala	Tyr	Gly	Ser	Ile	Thr	Phe	Gly	Asp	Val	Ala	Ile	Asp	Phe	Ser
1				5					10					15	
His	Gln	Glu	Trp	Glu	Tyr	Leu	Ser	Leu	Val	Gln	Lys	Thr	Leu	Tyr	Gln
			20					25					30		
Glu	Val	Met	Met	Glu	Asn	Tyr	Asp	Asn	Leu	Val	Ser	Leu	Ala	Gly	His
		35					40						45		
Ser	Val	Ser	Lys	Pro	Asp	Leu	Ile	Thr	Leu	Leu	Glu	Gln	Gly	Lys	Glu
	50					55					60				
Pro	Trp	Met	Ile	Val	Arg	Glu	Glu	Thr	Arg	Gly	Glu	Cys	Thr	Asp	Leu
	65				70					75				80	
Asp	Ser	Arg	Cys	Glu	Ile	Ile	Ser	Asp	Gly	Lys	Met	Gln	Leu	Tyr	Arg
			85						90					95	
Lys	His	Ser	Cys	Val	Thr	Leu	His	Gln	Arg	Ile	His	Asn	Gly	Gln	Lys
			100					105					110		
Pro	Tyr	Glu	Cys	Lys	Gln	Cys	Gln	Lys	Ser	Phe	Ser	His	Leu	Thr	Glu
		115					120						125		
Leu	Met	Val	His	Gln	Thr	Ile	His	Thr	Ser	Glu	Glu	Pro	Asp	Gln	Cys
		130					135					140			
Glu	Lys	Phe	Arg	Lys	Ala	Phe	Ser	His	Leu	Thr	Asp	Leu	Arg	Lys	His
	145				150					155				160	
Gln	Lys	Ile	Asn	Ala	Arg	Glu	Lys	Pro	Tyr	Glu	Cys	Glu	Glu	Cys	Gly
			165						170					175	
Lys	Val	Phe	Ser	Tyr	Pro	Ala	Asn	Leu	Ala	Gln	His	Gly	Lys	Val	His
			180						185				190		
Val	Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Glu	Cys	Gly	Glu	Ala	Phe	Arg	Thr
		195						200					205		
Ser	Arg	Gln	Leu	Thr	Val	His	His	Arg	Phe	His	Tyr	Gly	Glu	Lys	Pro
		210					215					220			

Tyr Glu Cys Lys Glu Cys Gly Lys Ala Phe Ser Val Tyr Gly Arg Leu
225 230 235 240
Ser Arg His Gln Ser Ile His Thr Gly Glu Lys Pro Phe Glu Cys Asn
245 250 255
Lys Cys Gly Lys Ser Phe Arg Leu Lys Ala Gly Leu Lys Val His Gln
260 265 270
Ser Ile His Thr Gly Glu Lys Pro His Glu Cys Lys Glu Cys Gly Lys
275 280 285
Ala Phe Arg Gln Phe Ser His Leu Val Gly His Lys Arg Ile His Thr
290 295 300
Gly Glu Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys Gly Phe Thr Cys
305 310 315 320
Arg Tyr Gln Leu Thr Met His Gln Arg Ile Tyr Ser Gly Glu Lys His
325 330 335
Tyr Glu Cys Lys Glu Asn Gly Glu Ala Phe Ser Ser Gly His Gln Leu
340 345 350
Thr Ala Pro His Thr Phe Glu Ser Val Glu Lys Pro Tyr Lys Cys Glu
355 360 365
Glu Cys Gly Lys Ala Phe Ser Val His Gly Arg Leu Thr Arg His Gln
370 375 380
Gly Ile His Ser Gly Lys Lys Pro Tyr Glu Cys Asn Lys Cys Gly Lys
385 390 395 400
Ser Phe Arg Leu Asn Ser Ser Leu Lys Ile His Gln Asn Ile His Thr
405 410 415
Gly Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala Phe Ser Gln
420 425 430
Arg Ala His Leu Ala His His Asn Arg Ile His Thr Gly Tyr Lys Pro
435 440 445
Phe Glu Cys Lys Glu Cys Gly Lys Ser Phe Arg Cys Ala Ser Tyr Leu

450 455 460
Val Ile His Glu Arg Ile His Thr Gly Glu Lys Pro Tyr Val Cys Gln
465 470 475 480
Glu Cys Gly Lys Gly Phe Ser Tyr Ser His Lys Leu Thr Ile His Arg
 485 490 495
Arg Val His Thr Gly Glu Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys
 500 505 510
Ala Phe Ser Val Ser Gly Gln Leu Thr Gln His Leu Ser Ile His Ser
 515 520 525
Gly Lys Arg Pro Phe Glu Cys Asn Lys Cys Gly Lys Ser Phe Arg Phe
 530 535 540
Ile Ser Val Leu Lys Ala His Gln Asn Ile His Ser Ala Glu Lys Pro
545 550 555 560
Tyr Glu Cys Lys Glu Cys Gly Lys Ala Phe Arg His Ala Thr Ser Leu
 565 570 575
Ile Tyr His Asp Arg Thr His Ala Gly Glu Lys Ser Tyr Glu Cys Lys
 580 585 590
Glu Cys Gly Glu Thr Phe Ser His Ala Ser His Leu Ile Ile His Glu
 595 600 605
Arg Ile His Thr Ser Asp Lys Pro Tyr Glu Cys Lys Arg Cys Gly Lys
 610 615 620
Ala Phe His Cys Ala Ser Tyr Leu Val Arg His Glu Ser Val His Ala
625 630 635 640
Asp Gly Asn Pro Tyr Met Cys Glu Glu Cys Gly Lys Ala Phe Asn Ser
 645 650 655
Ser His Glu Leu Ser Ile His His Arg Val His Thr Gly Glu Lys Pro
 660 665 670
Phe Lys Cys Asn Lys Cys Arg Arg Ser Phe Arg Leu Arg Ser Ile Leu
 675 680 685

Glu Val His Gln Arg Ile His Ile

690

695

<210> 65

<211> 669

<212> PRT

<213> Homo sapiens

<400> 65

Met Ala Glu His Arg Ser Met Asp Gly Arg Met Glu Ala Ala Thr Arg

1

5

10

15

Gly Gly Ser His Leu Gln Ile Ala Trp Ala Cys Gly Ser Pro Glu Ala

20

25

30

Leu Pro Pro Glu Gly Met Ala Ala Gln Thr His Ser Ala Gln Arg Cys

35

40

45

Leu Gln Thr Gly Pro Gly Cys Ser Pro Asp Ala Pro Gln His Gly Pro

50

55

60

Pro Ser Ala Pro Pro Pro Pro Pro Lys Glu Gly His Gln Glu Gly Leu

65

70

75

80

Val Glu Leu Pro Ala Ser Phe Arg Glu Leu Leu Thr Phe Phe Cys Thr

85

90

95

Asn Ala Thr Ile His Gly Ala Ile Arg Leu Val Cys Ser Arg Gly Asn

100

105

110

Arg Leu Lys Thr Thr Ser Trp Gly Leu Leu Ser Leu Gly Ala Leu Val

115

120

125

Ala Leu Cys Trp Gln Leu Gly Leu Leu Phe Glu Arg His Trp His Arg

130

135

140

Pro Val Leu Met Ala Val Ser Val His Ser Glu Arg Lys Leu Leu Pro

出証特 2 0 0 4 - 3 0 5 9 6 6 3

Ser Ile Arg Glu Asp Glu Val His Arg Leu Gly Ser Pro Tyr Gly His
385 390 395 400
Cys Thr Ala Gly Gly Glu Gly Val Glu Val Glu Leu Leu His Asn Thr
405 410 415
Ser Tyr Thr Arg Gln Ala Cys Leu Val Ser Cys Phe Gln Gln Leu Met
420 425 430
Val Glu Thr Cys Ser Cys Gly Tyr Tyr Leu His Pro Leu Pro Ala Gly
435 440 445
Ala Glu Tyr Cys Ser Ser Ala Arg His Pro Ala Trp Gly His Cys Phe
450 455 460
Tyr Arg Leu Tyr Gln Asp Leu Glu Thr His Arg Leu Pro Cys Thr Ser
465 470 475 480
Arg Cys Pro Arg Pro Cys Arg Glu Ser Ala Phe Lys Leu Ser Thr Gly
485 490 495
Thr Ser Arg Trp Pro Ser Ala Lys Ser Ala Gly Trp Thr Leu Ala Thr
500 505 510
Leu Gly Glu Gln Gly Leu Pro His Gln Ser His Arg Gln Arg Ser Ser
515 520 525
Leu Ala Lys Ile Asn Ile Val Tyr Gln Glu Leu Asn Tyr Arg Ser Val
530 535 540
Glu Glu Ala Pro Val Tyr Ser Val Pro Gln Leu Leu Ser Ala Met Gly
545 550 555 560
Ser Leu Cys Ser Leu Trp Phe Gly Ala Ser Val Leu Ser Leu Leu Glu
565 570 575
Leu Leu Glu Leu Leu Leu Asp Ala Ser Ala Leu Thr Leu Val Leu Gly
580 585 590
Gly Arg Arg Leu Arg Arg Ala Trp Phe Ser Trp Pro Arg Ala Ser Pro
595 600 605
Ala Ser Gly Ala Ser Ser Ile Lys Pro Glu Ala Ser Gln Met Pro Pro

610	615	620
Pro Ala Gly Gly Thr Ser Asp Asp Pro Glu Pro Ser Gly Pro His Leu		
625	630	635
Pro Arg Val Met Leu Pro Gly Val Leu Ala Gly Val Ser Ala Glu Glu		640
	645	650
Gly Arg Ala Gly Pro Gln Pro Leu Glu Thr Leu Asp Thr		655
660	665	

<210> 66

<211> 695

<212> PRT

<213> Homo sapiens

<400> 66

Met Ala Tyr Gly Ser Ile Thr Phe Gly Asp Val Ala Ile Asp Phe Ser		
1	5	10
His Gln Glu Trp Glu Tyr Leu Ser Leu Val Gln Lys Thr Leu Tyr Gln		
	20	25
Glu Val Met Met Glu Asn Tyr Asp Asn Leu Val Ser Leu Gly His Ser		
	35	40
Val Ser Lys Pro Asp Leu Ile Thr Leu Leu Glu Gln Gly Lys Glu Pro		
50	55	60
Trp Met Ile Val Arg Glu Glu Thr Arg Gly Glu Cys Thr Asp Leu Asp		
65	70	75
Ser Arg Cys Glu Ile Ile Ser Asp Gly Lys Met Gln Leu Tyr Arg Lys		
	85	90
His Ser Cys Val Thr Leu His Gln Arg Ile His Asn Gly Gln Lys Pro		
100	105	110

Tyr	Glu	Cys	Lys	Gln	Cys	Gln	Lys	Ser	Phe	Ser	His	Leu	Thr	Glu	Leu
115				120				125							
Met	Val	His	Gln	Thr	Ile	His	Thr	Ser	Glu	Glu	Pro	Asp	Gln	Cys	Glu
130				135				140							
Lys	Phe	Arg	Lys	Ala	Phe	Ser	His	Leu	Thr	Asp	Leu	Arg	Lys	His	Gln
145				150				155				160			
Lys	Ile	Asn	Ala	Arg	Glu	Lys	Pro	Tyr	Glu	Cys	Glu	Glu	Cys	Gly	Lys
				165				170				175			
Val	Phe	Ser	Tyr	Pro	Ala	Asn	Leu	Ala	Gln	His	Gly	Lys	Val	His	Val
				180				185				190			
Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Glu	Cys	Gly	Glu	Ala	Phe	Arg	Thr	Ser
195				200				205							
Arg	Gln	Leu	Thr	Val	His	His	Arg	Phe	His	Tyr	Gly	Glu	Lys	Pro	Tyr
210				215				220							
Glu	Cys	Lys	Glu	Cys	Gly	Lys	Ala	Phe	Ser	Val	Tyr	Gly	Arg	Leu	Ser
225				230				235				240			
Arg	His	Gln	Ser	Ile	His	Thr	Gly	Glu	Lys	Pro	Phe	Glu	Cys	Asn	Lys
				245				250				255			
Cys	Gly	Lys	Ser	Phe	Arg	Leu	Lys	Ala	Gly	Leu	Lys	Val	His	Gln	Ser
260				265				270							
Ile	His	Thr	Gly	Glu	Lys	Pro	His	Glu	Cys	Lys	Glu	Cys	Gly	Lys	Ala
275				280				285							
Phe	Arg	Gln	Phe	Ser	His	Leu	Val	Gly	His	Lys	Arg	Ile	His	Thr	Gly
290				295				300							
Glu	Lys	Pro	Tyr	Glu	Cys	Lys	Glu	Cys	Gly	Lys	Gly	Phe	Thr	Cys	Arg
305				310				315				320			
Tyr	Gln	Leu	Thr	Met	His	Gln	Arg	Ile	Tyr	Ser	Gly	Glu	Lys	His	Tyr
				325				330				335			
Glu	Cys	Lys	Glu	Asn	Gly	Glu	Ala	Phe	Ser	Ser	Gly	His	Gln	Leu	Thr

340	345	350
Ala Pro His Thr Phe Glu Ser Val Glu Lys Pro Tyr Lys Cys Glu Glu		
355	360	365
Cys Gly Lys Ala Phe Ser Val His Gly Arg Leu Thr Arg His Gln Gly		
370	375	380
Ile His Ser Gly Lys Lys Pro Tyr Glu Cys Asn Lys Cys Gly Lys Ser		
385	390	395
Phe Arg Leu Asn Ser Ser Leu Lys Ile His Gln Asn Ile His Thr Gly		
405	410	415
Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala Phe Ser Gln Arg		
420	425	430
Ala His Leu Ala His His Asn Arg Ile His Thr Gly Tyr Lys Pro Phe		
435	440	445
Glu Cys Lys Glu Cys Gly Lys Ser Phe Arg Cys Ala Ser Tyr Leu Val		
450	455	460
Ile His Glu Arg Ile His Thr Gly Glu Lys Pro Tyr Val Cys Gln Glu		
465	470	475
Cys Gly Lys Gly Phe Ser Tyr Ser His Lys Leu Thr Ile His Arg Arg		
485	490	495
Val His Thr Gly Glu Lys Pro Tyr Glu Cys Lys Glu Cys Gly Lys Ala		
500	505	510
Phe Ser Val Ser Gly Gln Leu Thr Gln His Leu Ser Ile His Ser Gly		
515	520	525
Lys Arg Pro Phe Glu Cys Asn Lys Cys Gly Lys Ser Phe Arg Phe Ile		
530	535	540
Ser Val Leu Lys Ala His Gln Asn Ile His Ser Ala Glu Lys Pro Tyr		
545	550	555
Glu Cys Lys Glu Cys Gly Lys Ala Phe Arg His Ala Thr Ser Leu Ile		
565	570	575

Tyr His Asp Arg Thr His Ala Gly Glu Lys Ser Tyr Glu Cys Lys Glu
 580 585 590
 Cys Gly Glu Thr Phe Ser His Ala Ser His Leu Ile Ile His Glu Arg
 595 600 605
 Ile His Thr Ser Asp Lys Pro Tyr Glu Cys Lys Arg Cys Gly Lys Ala
 610 615 620
 Phe His Cys Ala Ser Tyr Leu Val Arg His Glu Ser Val His Ala Asp
 625 630 635 640
 Gly Asn Pro Tyr Met Cys Glu Glu Cys Gly Lys Ala Phe Asn Ser Ser
 645 650 655
 His Glu Leu Ser Ile His His Arg Val His Thr Gly Glu Lys Pro Phe
 660 665 670
 Lys Cys Asn Lys Cys Arg Arg Ser Phe Arg Leu Arg Ser Ile Leu Glu
 675 680 685
 Val His Gln Arg Ile His Ile
 690 695

<210> 67

<211> 173

<212> PRT

<213> Homo sapiens

<400> 67

Met Gly Leu Arg Gln Gly Asn Gly Thr Val Ser Leu Gly Ile Leu Ala
 1 5 10 15
 Thr Asp Pro Ala Pro Pro Thr Glu Pro Ser Ser Pro Thr Pro Cys Leu
 20 25 30
 Arg Asp Gly Arg Val Ser Glu Asn Thr Val Cys Leu Ile Arg Lys Leu

35 40 45
 Leu Val Leu Asp Pro Gln Gln Arg Leu Ala Ala Ala Asp Val Leu Glu
 50 55 60
 Ala Leu Ser Ala Ile Ile Ala Ser Trp Gln Ser Leu Ser Ser Leu Ser
 65 70 75 80
 Gly Pro Leu Gln Val Val Pro Asp Ile Asp Asp Gln Met Ser Asn Ala
 85 90 95
 Asp Ser Ser Gln Glu Ala Lys Val Thr Glu Glu Cys Ser Gln Tyr Glu
 100 105 110
 Phe Glu Asn Tyr Met Arg Gln Gln Leu Leu Leu Ala Glu Glu Lys Ser
 115 120 125
 Ser Ile His Asp Ala Arg Ser Trp Val Pro Lys Arg Gln Phe Gly Ser
 130 135 140
 Ala Pro Pro Val Arg Arg Leu Gly His Asp Ala Gln Pro Met Thr Ser
 145 150 155 160
 Leu Asp Thr Ala Ile Leu Ala Gln Arg Tyr Leu Arg Lys
 165 170

<210> 68

<211> 867

<212> PRT

<213> Homo sapiens

<400> 68

Met Arg Gly Phe Asn Leu Leu Leu Phe Trp Gly Cys Cys Val Met His
 1 5 10 15
 Ser Trp Glu Gly His Ile Arg Pro Thr Arg Lys Pro Asn Thr Lys Gly
 20 25 30

Asn	Asn	Cys	Arg	Asp	Ser	Thr	Leu	Cys	Pro	Ala	Tyr	Ala	Thr	Cys	Thr
		35					40					45			
Asn	Thr	Val	Asp	Ser	Tyr	Tyr	Cys	Ala	Cys	Lys	Gln	Gly	Phe	Leu	Ser
		50					55					60			
Ser	Asn	Gly	Gln	Asn	His	Phe	Lys	Asp	Pro	Gly	Val	Arg	Cys	Lys	Asp
		65				70				75					80
Ile	Asn	Glu	Cys	Leu	Thr	Ser	Ser	Val	Cys	Pro	Glu	His	Ser	Asp	Cys
				85						90					95
Val	Asn	Ser	Met	Gly	Ser	Tyr	Ser	Cys	Ser	Cys	Gln	Val	Gly	Phe	Ile
			100						105					110	
Ser	Arg	Asn	Ser	Thr	Cys	Glu	Asp	Val	Asp	Glu	Cys	Ala	Asp	Pro	Arg
			115					120						125	
Ala	Cys	Pro	Glu	His	Ala	Thr	Cys	Asn	Asn	Thr	Val	Gly	Asn	Tyr	Ser
		130					135					140			
Cys	Phe	Cys	Asn	Pro	Gly	Phe	Glu	Ser	Ser	Ser	Gly	His	Leu	Ser	Phe
		145					150				155				160
Gln	Gly	Leu	Lys	Ala	Ser	Cys	Glu	Asp	Ile	Asp	Glu	Cys	Thr	Glu	Met
				165					170						175
Cys	Pro	Ile	Asn	Ser	Thr	Cys	Thr	Asn	Thr	Pro	Gly	Ser	Tyr	Phe	Cys
			180						185					190	
Thr	Cys	His	Pro	Gly	Phe	Ala	Pro	Ser	Asn	Gly	Gln	Leu	Asn	Phe	Thr
			195						200					205	
Asp	Gln	Gly	Val	Glu	Cys	Arg	Asp	Ile	Asp	Glu	Cys	Arg	Gln	Asp	Pro
		210						215					220		
Ser	Thr	Cys	Gly	Pro	Asn	Ser	Ile	Cys	Thr	Asn	Ala	Leu	Gly	Ser	Tyr
		225					230				235				240
Ser	Cys	Gly	Cys	Ile	Ala	Gly	Phe	His	Pro	Asn	Pro	Glu	Gly	Ser	Gln
				245						250					255
Lys	Asp	Gly	Asn	Phe	Ser	Cys	Gln	Arg	Val	Leu	Phe	Lys	Cys	Lys	Glu

260 265 270
Asp Val Ile Pro Asp Asn Lys Gln Ile Gln Gln Cys Gln Glu Gly Thr
275 280 285
Ala Val Lys Pro Ala Tyr Val Ser Phe Cys Ala Gln Ile Asn Asn Ile
290 295 300
Phe Ser Val Leu Asp Lys Val Cys Glu Asn Lys Thr Thr Val Val Ser
305 310 315 320
Leu Lys Asn Thr Thr Glu Ser Phe Val Pro Val Leu Lys Gln Ile Ser
325 330 335
Thr Trp Thr Lys Phe Thr Lys Glu Glu Thr Ser Ser Leu Ala Thr Val
340 345 350
Phe Leu Glu Ser Val Glu Ser Met Thr Leu Ala Ser Phe Trp Lys Pro
355 360 365
Ser Ala Asn Val Thr Pro Ala Val Arg Thr Glu Tyr Leu Asp Ile Glu
370 375 380
Ser Lys Val Ile Asn Lys Glu Cys Ser Glu Glu Asn Val Thr Leu Asp
385 390 395 400
Leu Val Ala Lys Gly Asp Lys Met Lys Ile Gly Cys Ser Thr Ile Glu
405 410 415
Glu Ser Glu Ser Thr Glu Thr Thr Gly Val Ala Phe Val Ser Phe Val
420 425 430
Gly Met Glu Ser Val Leu Asn Glu Arg Phe Phe Gln Asp His Gln Ala
435 440 445
Pro Leu Thr Thr Ser Glu Ile Lys Leu Lys Met Asn Ser Arg Val Val
450 455 460
Gly Gly Ile Met Thr Gly Glu Lys Lys Asp Gly Phe Ser Asp Pro Ile
465 470 475 480
Ile Tyr Thr Leu Glu Asn Val Gln Pro Lys Gln Lys Phe Glu Arg Pro
485 490 495

Ile Cys Val Ser Trp Ser Thr Asp Val Lys Gly Gly Arg Trp Thr Ser
500 505 510
Phe Gly Cys Val Ile Leu Glu Ala Ser Glu Thr Tyr Thr Ile Cys Ser
515 520 525
Cys Asn Gln Met Ala Asn Leu Ala Val Ile Met Ala Ser Gly Glu Leu
530 535 540
Thr Met Asp Phe Ser Leu Tyr Ile Ile Ser His Val Gly Ile Ile Ile
545 550 555 560
Ser Leu Val Cys Leu Val Leu Ala Ile Ala Thr Phe Leu Leu Cys Arg
565 570 575
Ser Ile Arg Asn His Asn Thr Tyr Leu His Leu His Leu Cys Val Cys
580 585 590
Leu Leu Leu Ala Lys Thr Leu Phe Leu Ala Gly Ile His Lys Thr Asp
595 600 605
Asn Lys Met Gly Cys Ala Ile Ile Ala Gly Phe Leu His Tyr Leu Phe
610 615 620
Leu Ala Cys Phe Phe Trp Met Leu Val Glu Ala Val Ile Leu Phe Leu
625 630 635 640
Met Val Arg Asn Leu Lys Val Val Asn Tyr Phe Ser Ser Arg Asn Ile
645 650 655
Lys Met Leu His Ile Cys Ala Phe Gly Tyr Gly Leu Pro Met Leu Val
660 665 670
Val Val Ile Ser Ala Ser Val Gln Pro Gln Gly Tyr Gly Met His Asn
675 680 685
Arg Cys Trp Leu Asn Thr Glu Thr Gly Phe Ile Trp Ser Phe Leu Gly
690 695 700
Pro Val Cys Thr Val Ile Val Val Ser Lys Tyr Tyr Asn Ser Leu Ala
705 710 715 720
Lys Cys Val Leu Lys Glu Glu Gln Gly Asp Leu Arg Asp Leu Glu Phe

725	730	735
Pro Gly Thr Cys Ala Ala Glu Arg Ile Asn Ser Leu Leu Leu Thr Trp		
740	745	750
Thr Leu Trp Ile Leu Arg Gln Arg Leu Ser Ser Val Asn Ala Glu Val		
755	760	765
Ser Thr Leu Lys Asp Thr Arg Leu Leu Thr Phe Lys Ala Phe Ala Gln		
770	775	780
Leu Phe Ile Leu Gly Cys Ser Trp Val Leu Gly Ile Phe Gln Ile Gly		
785	790	795
800		
Pro Val Ala Gly Val Met Ala Tyr Leu Phe Thr Ile Ile Asn Ser Leu		
805	810	815
Gln Gly Ala Phe Ile Phe Leu Ile His Cys Leu Leu Asn Gly Gln Val		
820	825	830
Arg Glu Glu Tyr Lys Arg Trp Ile Thr Gly Lys Thr Lys Pro Ser Ser		
835	840	845
Gln Ser Gln Thr Ser Arg Ile Leu Leu Ser Ser Met Pro Ser Ala Ser		
850	855	860
Lys Thr Gly		
865		

<210> 69

<211> 317

<212> PRT

<213> Homo sapiens

<400> 69

Met Arg Leu Ala Leu His Ser Trp Val Ala Leu Leu Phe Leu Thr Leu

1

5

10

15

Ser Ser Tyr Thr Ser Thr Gln Val Leu Ile Glu Thr Leu Cys Ala Leu
 20 25 30
 Gly Ala Gln Cys Arg Trp Ser Ala Cys Asn Ile Tyr Ser Thr Gln Asn
 35 40 45
 Glu Val Ala Ala Ala Leu Ala Glu Ala Gly Val Ala Val Phe Ala Trp
 50 55 60
 Lys Gly Glu Ser Glu Asp Asp Phe Trp Trp Cys Ile Asp Arg Cys Val
 65 70 75 80
 Asn Met Asp Gly Trp Gln Ala Asn Met Ile Leu Asp Asp Gly Gly Asp
 85 90 95
 Leu Thr His Trp Val Tyr Lys Lys Tyr Pro Asn Val Phe Lys Lys Ile
 100 105 110
 Arg Gly Ile Val Glu Glu Ser Val Thr Gly Val His Arg Leu Tyr Gln
 115 120 125
 Leu Ser Lys Ala Gly Lys Leu Cys Val Pro Ala Met Asn Val Asn Asp
 130 135 140
 Ser Val Thr Lys Gln Lys Phe Asp Asn Leu Tyr Cys Cys Arg Glu Ser
 145 150 155 160
 Ile Leu Asp Gly Leu Lys Arg Thr Thr Asp Val Met Phe Gly Gly Lys
 165 170 175
 Gln Val Val Val Cys Gly Tyr Gly Glu Val Gly Lys Gly Cys Cys Ala
 180 185 190
 Ala Leu Lys Ala Leu Gly Ala Ile Val Tyr Ile Thr Glu Ile Asp Pro
 195 200 205
 Ile Cys Ala Leu Gln Ala Cys Met Asp Gly Phe Arg Val Val Lys Leu
 210 215 220
 Asn Glu Val Ile Arg Gln Val Asp Val Val Ile Thr Cys Thr Gly Asn
 225 230 235 240
 Lys Asn Val Val Thr Arg Glu His Leu Asp Arg Met Lys Asn Ser Cys

245 250 255
 Ile Val Cys Asn Met Gly His Ser Asn Thr Glu Ile Asp Val Thr Ser
 260 265 270
 Leu Arg Thr Pro Glu Leu Thr Trp Glu Arg Val Arg Ser Gln Val Asp
 275 280 285
 His Val Ile Trp Pro Asp Gly Lys Arg Val Val Leu Leu Ala Glu Thr
 290 295 300
 Gln Met Asp Leu Gln Lys Ile Gly Lys Ser Leu Glu Ser
 305 310 315

<210> 70

<211> 501

<212> PRT

<213> Homo sapiens

<400> 70

Met Asp Gly Lys Val Ala Val Gln Glu Cys Gly Pro Pro Ala Val Ser
 1 5 10 15
 Trp Val Pro Glu Glu Gly Glu Lys Leu Asp Gln Glu Asp Glu Asp Gln
 20 25 30
 Val Lys Asp Arg Gly Gln Trp Thr Asn Lys Val Glu Phe Val Leu Ser
 35 40 45
 Val Ala Gly Glu Ile Ile Gly Leu Ala Ser Val Val Ile Glu Ser Tyr
 50 55 60
 Leu Asn Val Tyr Tyr Ile Ile Ile Leu Ala Trp Ala Leu Phe Tyr Leu
 65 70 75 80
 Phe Ser Ser Phe Thr Ser Glu Leu Pro Trp Thr Thr Cys Asn Asn Phe
 85 90 95

Trp Asn Thr Glu His Cys Thr Asp Phe Leu Asn His Ser Gly Ala Gly
100 105 110
Thr Val Thr Pro Phe Glu Asn Phe Thr Ser Pro Val Met Glu Phe Trp
115 120 125
Glu Arg Arg Val Leu Gly Ile Thr Ser Gly Ile His Asp Leu Gly Ser
130 135 140
Leu Arg Trp Glu Leu Ala Leu Cys Leu Leu Leu Ala Trp Val Ile Cys
145 150 155 160
Tyr Phe Cys Ile Trp Lys Gly Val Lys Ser Thr Gly Lys Val Val Tyr
165 170 175
Phe Thr Ala Thr Phe Pro Tyr Leu Met Leu Val Ile Leu Leu Ile Arg
180 185 190
Gly Val Thr Leu Pro Gly Ala Tyr Gln Gly Ile Ile Tyr Tyr Leu Lys
195 200 205
Pro Asp Leu Phe Arg Leu Lys Asp Pro Gln Val Trp Met Asp Ala Gly
210 215 220
Thr Gln Ile Phe Phe Ser Phe Ala Ile Cys Gln Gly Cys Leu Thr Ala
225 230 235 240
Leu Gly Ser Tyr Asn Lys Tyr His Asn Asn Cys Tyr Lys Asp Cys Ile
245 250 255
Ala Leu Cys Phe Leu Asn Ser Ala Thr Ser Phe Val Ala Gly Phe Val
260 265 270
Val Phe Ser Ile Leu Gly Phe Met Ser Gln Glu Gln Gly Val Pro Ile
275 280 285
Ser Glu Val Ala Glu Ser Gly Pro Gly Leu Ala Phe Ile Ala Phe Pro
290 295 300
Lys Ala Val Thr Met Met Pro Leu Ser Gln Leu Trp Ser Cys Leu Phe
305 310 315 320
Phe Ile Met Leu Ile Phe Leu Gly Leu Asp Ser Gln Phe Val Cys Val

325	330	335
Glu Cys Leu Val Thr Ala Ser Ile Asp Met Phe Pro Arg Gln Leu Arg		
340	345	350
Lys Ser Gly Arg Arg Glu Leu Leu Ile Leu Thr Ile Ala Val Met Cys		
355	360	365
Tyr Leu Ile Gly Leu Phe Leu Val Thr Glu Gly Gly Met Tyr Ile Phe		
370	375	380
Gln Leu Phe Asp Tyr Tyr Ala Ser Ser Gly Ile Cys Leu Leu Phe Leu		
385	390	395
Ser Leu Phe Glu Val Val Cys Ile Ser Trp Val Tyr Gly Ala Asp Arg		
405	410	415
Phe Tyr Asp Asn Ile Glu Asp Met Ile Gly Trp Phe Leu Ala Leu Ser		
420	425	430
Ser Met Val Cys Val Pro Leu Phe Val Val Ile Thr Leu Leu Lys Thr		
435	440	445
Arg Gly Pro Phe Arg Lys Arg Leu Arg Gln Leu Ile Thr Pro Asp Ser		
450	455	460
Ser Leu Pro Gln Pro Lys Gln His Pro Cys Leu Asp Gly Ser Ala Gly		
465	470	475
Arg Asn Phe Gly Pro Ser Pro Thr Arg Glu Gly Leu Ile Ala Gly Glu		
485	490	495
Lys Glu Thr His Leu		
500		

<210> 71

<211> 126

<212> PRT

<213> Homo sapiens

<400> 71

Met	Thr	Val	Ser	Arg	Ser	Cys	Ser	Lys	Leu	Pro	Met	His	Phe	Ser	Val
1				5					10					15	
Phe	Leu	Lys	Gln	Asn	Leu	Asn	Ser	Leu	Thr	Cys	Asn	Lys	Ala	Cys	Val
			20					25					30		
Ala	Ala	Leu	Leu	Ser	Trp	Leu	Gln	Gly	Phe	Ser	Ser	Val	Gly	Ala	Cys
			35				40					45			
Gly	Ile	Leu	Ser	Phe	Leu	Thr	Cys	Val	Trp	Glu	Gln	Gln	Tyr	Leu	Ser
	50					55					60				
His	Arg	Val	Ala	His	Gly	Ile	Ser	Gly	Ala	Val	His	Ala	Lys	Cys	Leu
65					70				75					80	
Ala	Trp	Cys	Leu	Ile	Val	Lys	Ala	Gly	Trp	Leu	Thr	Val	Phe	Thr	Leu
				85				90					95		
Asp	Leu	Gln	Lys	Trp	Ser	Trp	Pro	Ile	Pro	Gln	Lys	Thr	Ala	Glu	Arg
			100				105					110			
Thr	Pro	Gly	Lys	Met	Met	Thr	Tyr	Ser	Ser	Cys	Ser	Ser	Pro		
		115				120						125			

<210> 72

<211> 652

<212> PRT

<213> Homo sapiens

<400> 72

Met	Thr	Ser	Ala	Gln	Tyr	Lys	Ile	Lys	Ile	Leu	Ile	Glu	Gly	Leu	His
1				5					10				15		
His	Gly	Gln	Arg	Pro	Gly	Pro	Ala	Pro	Pro	Arg	Gln	Pro	Gly	Ser	Phe

20	25	30
Cys Trp Ala Leu Lys Ala Asp Gly Ile Met Trp Leu Ala Lys Ala Cys		
35	40	45
Trp Ser Ile Gln Ser Glu Met Pro Cys Ile Gln Ala Gln Tyr Gly Thr		
50	55	60
Pro Ala Pro Ser Pro Gly Pro Arg Asp His Leu Ala Ser Asp Pro Leu		
65	70	75
Thr Pro Glu Phe Ile Lys Pro Thr Met Asp Leu Ala Ser Pro Glu Ala		
85	90	95
Ala Pro Ala Ala Pro Thr Ala Leu Pro Ser Phe Ser Thr Phe Met Asp		
100	105	110
Gly Tyr Thr Gly Glu Phe Asp Thr Phe Leu Tyr Gln Leu Pro Gly Thr		
115	120	125
Val Gln Pro Cys Ser Ser Ala Ser Ser Ser Ala Ser Ser Thr Ser Ser		
130	135	140
Ser Ser Ala Thr Ser Pro Ala Ser Ala Ser Phe Lys Phe Glu Asp Phe		
145	150	155
Gln Val Tyr Gly Cys Tyr Pro Gly Pro Leu Ser Gly Pro Val Asp Glu		
165	170	175
Ala Leu Ser Ser Ser Gly Ser Asp Tyr Tyr Gly Ser Pro Cys Ser Ala		
180	185	190
Pro Ser Pro Ser Thr Pro Ser Phe Gln Pro Pro Gln Leu Ser Pro Trp		
195	200	205
Asp Gly Ser Phe Gly His Phe Ser Pro Ser Gln Thr Tyr Glu Gly Leu		
210	215	220
Arg Ala Trp Thr Glu Gln Leu Pro Lys Ala Ser Gly Pro Pro Gln Pro		
225	230	235
Pro Ala Phe Phe Ser Phe Ser Pro Pro Thr Gly Pro Ser Pro Ser Leu		
245	250	255

Ala Gln Ser Pro Leu Lys Leu Phe Pro Ser Gln Ala Thr His Gln Leu
 260 265 270
 Gly Glu Gly Glu Ser Tyr Ser Met Pro Thr Ala Phe Pro Gly Leu Ala
 275 280 285
 Pro Thr Ser Pro His Leu Glu Gly Ser Gly Ile Leu Asp Thr Pro Val
 290 295 300
 Thr Ser Thr Lys Ala Arg Ser Gly Ala Pro Gly Gly Ser Glu Gly Arg
 305 310 315 320
 Cys Ala Val Cys Gly Asp Asn Ala Ser Cys Gln His Tyr Gly Val Arg
 325 330 335
 Thr Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Gln Lys Asn
 340 345 350
 Ala Lys Tyr Ile Cys Leu Ala Asn Lys Asp Cys Pro Val Asp Lys Arg
 355 360 365
 Arg Arg Asn Arg Cys Gln Phe Cys Arg Phe Gln Lys Cys Leu Ala Val
 370 375 380
 Gly Met Val Lys Glu Val Val Arg Thr Asp Ser Leu Lys Gly Arg Arg
 385 390 395 400
 Gly Arg Leu Pro Ser Lys Pro Lys Gln Pro Pro Asp Ala Ser Pro Ala
 405 410 415
 Asn Leu Leu Thr Ser Leu Val Arg Ala His Leu Asp Ser Gly Pro Ser
 420 425 430
 Thr Ala Lys Leu Asp Tyr Ser Lys Phe Gln Glu Leu Val Leu Pro His
 435 440 445
 Phe Gly Lys Glu Asp Ala Gly Asp Val Gln Gln Phe Tyr Asp Leu Leu
 450 455 460
 Ser Gly Ser Leu Glu Val Ile Arg Lys Trp Ala Glu Lys Ile Pro Gly
 465 470 475 480
 Phe Ala Glu Leu Ser Pro Ala Asp Gln Asp Leu Leu Leu Glu Ser Ala

485 490 495
Phe Leu Glu Leu Phe Ile Leu Arg Leu Ala Tyr Arg Ser Lys Pro Gly
500 505 510
Glu Gly Lys Leu Ile Phe Cys Ser Gly Leu Val Leu His Arg Leu Gln
515 520 525
Cys Ala Arg Gly Phe Gly Asp Arg Ile Asp Ser Ile Leu Ala Phe Ser
530 535 540
Arg Ser Leu His Ser Leu Leu Val Asp Val Pro Ala Phe Ala Cys Leu
545 550 555 560
Ser Ala Leu Val Leu Ile Thr Asp Arg His Gly Leu Gln Glu Pro Arg
565 570 575
Arg Val Glu Glu Leu Gln Asn Arg Ile Ala Ser Cys Leu Lys Glu His
580 585 590
Val Ala Ala Val Ala Gly Glu Pro Gln Pro Ala Ser Cys Leu Ser Arg
595 600 605
Leu Leu Gly Lys Leu Pro Glu Leu Arg Thr Leu Cys Thr Gln Gly Leu
610 615 620
Gln Arg Ile Phe Tyr Leu Lys Leu Glu Asp Leu Val Pro Pro Pro Pro
625 630 635 640
Ile Ile Asp Lys Ile Phe Met Asp Thr Leu Pro Phe
645 650

<210> 73

<211> 528

<212> PRT

<213> Homo sapiens

<400> 73

Met Gly Pro Leu Gln Phe Arg Asp Val Ala Ile Glu Phe Ser Leu Glu
 1 5 10 15
 Glu Trp His Cys Leu Asp Thr Ala Gln Arg Asn Leu Tyr Arg Asn Val
 20 25 30
 Met Leu Glu Asn Tyr Ser Asn Leu Val Phe Leu Gly Ile Thr Val Ser
 35 40 45
 Lys Pro Asp Leu Ile Thr Cys Leu Glu Gln Gly Arg Lys Pro Leu Thr
 50 55 60
 Met Lys Arg Asn Glu Met Ile Ala Lys Pro Ser Val Met Cys Ser His
 65 70 75 80
 Phe Ala Gln Asp Leu Trp Pro Glu Gln Ser Met Lys Asp Ser Phe Gln
 85 90 95
 Lys Val Val Leu Arg Arg Tyr Glu Lys Cys Glu His Asp Asn Leu Gln
 100 105 110
 Leu Lys Lys Gly Cys Ile Ser Val Asp Glu Cys Lys Val His Lys Glu
 115 120 125
 Gly Tyr Asn Glu Leu Asn Gln Cys Leu Thr Thr Thr Pro Arg Lys Ile
 130 135 140
 Cys Gln Cys Asp Lys Tyr Val Lys Val Leu His Gln Phe Pro Asn Ser
 145 150 155 160
 Asn Gly Gln Lys Arg Gly His Thr Gly Lys Lys Pro Phe Lys Tyr Ile
 165 170 175
 Glu Cys Gly Lys Ala Phe Lys Gln Phe Ser Thr Leu Thr Thr His Lys
 180 185 190
 Lys Ile His Thr Gly Gly Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys
 195 200 205
 Ala Phe Asn His Ser Cys Ser Leu Thr Arg His Lys Lys Ile His Thr
 210 215 220
 Gly Glu Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Lys His

225	230	235	240
Ser Ser Thr Leu Thr Thr His Lys Arg Asn His Thr Gly Glu Lys Pro			
	245	250	255
Tyr Lys Cys Asp Lys Cys Gly Lys Ala Phe Met Ser Ser Ser Thr Leu			
	260	265	270
Ser Lys His Glu Ile Ile His Thr Glu Lys Lys Pro Tyr Lys Cys Glu			
	275	280	285
Glu Cys Gly Lys Ala Phe Asn Arg Ser Ser Thr Leu Thr Thr His Lys			
	290	295	300
Ile Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Glu Glu Cys Asp Lys			
305	310	315	320
Ala Phe Lys Tyr Ser Tyr Thr Leu Thr Thr His Lys Arg Ile His Thr			
	325	330	335
Glu Asp Lys Pro Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Lys Tyr			
	340	345	350
Ser Ser Thr Leu Thr Thr His Lys Arg Ile His Thr Gly Glu Lys Pro			
	355	360	365
Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Lys Arg Ser Ser Asp Leu			
	370	375	380
Thr Thr His Lys Ile Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Glu			
385	390	395	400
Glu Cys Gly Lys Ala Phe Lys Tyr Ser Ser Asn Leu Thr Thr His Lys			
	405	410	415
Lys Ile His Thr Gly Glu Arg Pro Tyr Lys Cys Glu Glu Cys Gly Lys			
	420	425	430
Ala Phe Asn Gln Ser Ser Ile Leu Thr Thr His Arg Arg Ile His Thr			
	435	440	445
Gly Glu Lys Phe Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Lys Cys			
	450	455	460

Ser Ser Asn Leu Thr Thr His Lys Lys Ile His Thr Gly Glu Arg Pro
 465 470 475 480
 Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Asn Gln Ser Ser Ile Leu
 485 490 495
 Thr Thr His Glu Arg Ile Ile Leu Glu Arg Asn Ser Thr Asn Val Lys
 500 505 510
 Asn Val Ala Lys Pro Ser Ser Gly Pro His Thr Leu Leu His Ile Arg
 515 520 525

<210> 74

<211> 435

<212> PRT

<213> Homo sapiens

<400> 74

Met Pro Thr Ala Ser Ser Leu Val His Leu Asp Leu Gly Pro Pro Leu
 1 5 10 15
 Lys Val Pro Cys Gly Pro Phe Leu Leu Thr Ala Val Gly Glu Gly Leu
 20 25 30
 Ala Cys His Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro
 35 40 45
 Thr Gln Cys Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val
 50 55 60
 Glu Glu Cys Arg Val Leu Gln Gly Tyr Glu Gly Arg Arg Arg Gly Trp
 65 70 75 80
 Leu Glu Gly Cys Met Gly Leu Leu Ser Asp Pro Leu Thr Thr Val Pro
 85 90 95
 Ser Leu Arg Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu Pro

100	105	110
Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe Gly		
115	120	125
Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro Pro		
130	135	140
Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser Tyr		
145	150	155
Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro Cys		
165	170	175
Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly Cys		
180	185	190
Pro Ala Glu Gln Arg Ala Asn Pro Leu Thr Ser Ile Val Ser Ala Val		
195	200	205
Val Gly Ile Leu Leu Val Val Val Leu Gly Val Val Phe Gly Ile Leu		
210	215	220
Ile Lys Arg Arg Gln Gln Lys Ile Arg Lys Tyr Thr Met Arg Arg Leu		
225	230	235
Leu Gln Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Ala Met		
245	250	255
Pro Asn Gln Ala Gln Met Arg Ile Leu Lys Glu Thr Glu Leu Arg Lys		
260	265	270
Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile		
275	280	285
Trp Ile Pro Asp Gly Glu Asn Val Lys Ile Pro Val Ala Ile Lys Val		
290	295	300
Leu Arg Glu Asn Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu		
305	310	315
Ala Tyr Val Met Ala Gly Val Gly Ser Pro Tyr Val Ser Arg Leu Leu		
325	330	335

Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Val Thr Gln Leu Met Pro

340

345

350

Tyr Gly Cys Leu Leu Asp His Val Arg Glu Asn Arg Gly Arg Leu Gly

355

360

365

Ser Gln Asp Leu Leu Asn Trp Cys Met Gln Ile Ala Lys Gly Met Ser

370

375

380

Tyr Leu Glu Asp Val Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn

385

390

395

400

Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly Leu

405

410

415

Ala Arg Leu Leu Asp Ile Asp Glu Thr Glu Tyr His Ala Asp Gly Gly

420

425

430

Lys Val Arg

435

<210> 75

<211> 242

<212> PRT

<213> Homo sapiens

<400> 75

Met Gln Gly Glu Gln Gly Arg Arg Leu Phe Arg Leu Leu Leu Asn Asp

1

5

10

15

Leu Pro Gly Arg Asp Arg Leu Pro Cys Ser Tyr Pro Gln Pro His Val

20

25

30

Glu Gln Ala Ser Thr Ser Ala Ser Pro Arg Leu Leu Ala Val Asn Ile

35

40

45

Pro Glu Asp Thr Leu Leu Ser Tyr Pro Glu Ser Ser Leu Pro Leu Gln

50	55	60
Arg His Cys Ala Phe His Arg Phe Ser Lys Ala Lys Phe Glu Ser Leu		
65	70	75
Phe Arg Thr Tyr Asp Lys Asp Ile Thr Phe Gln Tyr Phe Lys Ser Phe		
85	90	95
Lys Arg Val Arg Ile Asn Phe Ser Asn Pro Phe Ser Ala Ala Asp Ala		
100	105	110
Arg Leu Gln Leu His Lys Thr Glu Phe Leu Gly Lys Glu Met Lys Leu		
115	120	125
Tyr Phe Ala Gln Thr Leu His Ile Gly Ser Ser His Leu Ala Pro Pro		
130	135	140
Asn Pro Asp Lys Gln Phe Leu Ile Ser Pro Pro Ala Ser Pro Pro Val		
145	150	155
Gly Trp Lys Gln Val Glu Asp Ala Thr Pro Val Ile Asn Tyr Asp Leu		
165	170	175
Leu Tyr Ala Ile Thr Lys Leu Gly Pro Gly Glu Lys Tyr Glu Leu His		
180	185	190
Ala Ala Thr Asp Thr Thr Pro Ser Val Val Val His Val Cys Glu Ser		
195	200	205
Asp Gln Glu Lys Glu Glu Glu Glu Glu Met Glu Arg Met Arg Arg Pro		
210	215	220
Lys Pro Lys Ile Ile Gln Thr Arg Arg Pro Glu Tyr Thr Pro Ile His		
225	230	235
Leu Ser		240

<210> 76

<211> 110

<212> PRT

<213> Homo sapiens

<400> 76

Met	Ala	Leu	Phe	Phe	Gly	Leu	Phe	Phe	Val	Phe	Glu	Thr	Glu	Ser	Arg
1			5						10					15	
Ser	Val	Thr	Lys	Ala	Gly	Val	Gln	Trp	Arg	Asp	Leu	Ser	Ser	Leu	Gln
			20					25						30	
Pro	Pro	Pro	Pro	Glu	Phe	Lys	Gln	Leu	Leu	Cys	Pro	Ser	Leu	Leu	Ser
			35					40						45	
Cys	Trp	Asp	Tyr	Arg	His	Thr	Thr	Leu	His	Pro	Ala	Lys	Ile	Phe	Val
			50					55						60	
Phe	Leu	Val	Glu	Thr	Glu	Phe	Arg	His	Val	Gly	Gln	Ala	Gly	Leu	Glu
			65					70						75	
Leu	Leu	Thr	Pro	Ser	Asp	Leu	Pro	Ala	Ala	Val	Ser	Gln	Ser	Ile	Gly
								85						90	
Thr	Thr	Gly	Met	Ser	His	His	Ala	Trp	Leu	Lys	Ala	Ser	Ile		
			100						105					110	

<210> 77

<211> 471

<212> PRT

<213> Homo sapiens

<400> 77

Met	Asp	Ser	Ala	Pro	Leu	Phe	Pro	Arg	Pro	His	Leu	Phe	Gln	Asn	Leu
1					5					10				15	
Leu	Leu	Phe	Leu	Trp	Ala	Leu	Leu	Asn	Cys	Gly	Leu	Gly	Val	Ser	Ala

20	25	30
Gln Gly Pro Gly Glu Trp Thr Pro Trp Val Ser Trp Thr Arg Cys Ser		
35	40	45
Ser Ser Cys Gly Arg Gly Val Ser Val Arg Ser Arg Arg Cys Leu Arg		
50	55	60
Leu Pro Gly Glu Glu Pro Cys Trp Gly Asp Ser His Glu Tyr Arg Leu		
65	70	75
Cys Gln Leu Pro Asp Cys Pro Pro Gly Ala Val Pro Phe Arg Asp Leu		
85	90	95
Gln Cys Ala Leu Tyr Asn Gly Arg Pro Val Leu Gly Thr Gln Lys Thr		
100	105	110
Tyr Gln Trp Val Pro Phe His Gly Ala Pro Asn Gln Cys Asp Leu Asn		
115	120	125
Cys Leu Ala Glu Gly His Ala Phe Tyr His Ser Phe Gly Arg Val Leu		
130	135	140
Asp Gly Thr Ala Cys Ser Pro Gly Ala Gln Gly Val Cys Val Ala Gly		
145	150	155
Arg Cys Leu Ser Ala Gly Cys Asp Gly Leu Leu Gly Ser Gly Ala Leu		
165	170	175
Glu Asp Arg Cys Gly Arg Cys Gly Gly Ala Asp Asp Ser Cys Leu Phe		
180	185	190
Val Gln Arg Val Phe Arg Asp Ala Gly Ala Phe Ala Gly Tyr Trp Asn		
195	200	205
Val Thr Leu Ile Pro Glu Gly Ala Arg His Ile Arg Val Glu His Arg		
210	215	220
Ser Arg Asn His Leu Ala Leu Met Gly Gly Asp Gly Arg Tyr Val Leu		
225	230	235
Asn Gly His Trp Val Val Ser Pro Pro Gly Thr Tyr Glu Ala Ala Gly		
245	250	255

Thr His Val Val Tyr Thr Arg Asp Thr Gly Pro Gln Glu Thr Leu Gln
260 265 270

Ala Ala Gly Pro Thr Ser His Asp Leu Leu Leu Gln Val Leu Leu Gln
275 280 285

Glu Pro Asn Pro Gly Ile Glu Phe Glu Phe Trp Leu Pro Arg Glu Arg
290 295 300

Tyr Ser Pro Phe Gln Ala Arg Val Gln Ala Leu Gly Trp Pro Leu Arg
305 310 315 320

Gln Pro Gln Pro Arg Gly Val Glu Pro Gln Pro Pro Ala Ala Pro Ala
325 330 335

Val Thr Pro Ala Gln Thr Pro Thr Leu Ala Pro Asp Pro Cys Pro Pro
340 345 350

Cys Pro Asp Thr Arg Gly Arg Ala His Arg Leu Leu His Tyr Cys Gly
355 360 365

Ser Asp Phe Val Phe Gln Ala Arg Val Leu Gly His His His Gln Ala
370 375 380

Gln Glu Thr Arg Tyr Glu Val Arg Ile Gln Leu Val Tyr Lys Asn Arg
385 390 395 400

Ser Pro Leu Arg Ala Arg Glu Tyr Val Trp Ala Pro Gly His Cys Pro
405 410 415

Cys Pro Met Leu Ala Pro His Arg Asp Tyr Leu Met Ala Val Gln Arg
420 425 430

Leu Val Ser Pro Asp Gly Thr Gln Asp Gln Leu Leu Leu Pro His Ala
435 440 445

Gly Tyr Ala Arg Pro Trp Ser Pro Ala Glu Asp Ser Arg Ile Arg Leu
450 455 460

Thr Ala Arg Arg Cys Pro Gly
465 470

<210> 78

<211> 149

<212> PRT

<213> Homo sapiens

<400> 78

Met Val Phe Pro Trp Ser Trp Ala Ser Arg Trp Pro Arg Leu Phe Ser

1

5

10

15

Asp Cys Pro Gly Gln Thr Pro Pro Cys Pro Ala Gly Gln Trp Pro Ala

20

25

30

Ser Met Pro Val Pro Leu Gly Val Leu Phe Cys Trp Arg Ala Pro Leu

35

40

45

Asp Asp Gln Leu Leu Gly Ser Ser Thr Glu Val Phe Leu Thr Met Ser

50

55

60

Ser Arg Ser Cys Val Cys Pro Leu Gly Ser Arg Val Phe Ile Gly Pro

65

70

75

80

Gly Leu Gly His Gly Arg Pro Arg Trp Ser Trp Lys Met Val Thr Phe

85

90

95

Gly His Glu Ser Arg Ser Ala Tyr Pro His Leu Asp Leu Trp Gly Trp

100

105

110

Ser Pro Ser Gln Gly Pro Thr Phe Leu Tyr Pro Ala Leu Pro Leu Pro

115

120

125

Leu Ser Val Ser Phe Tyr Gly Thr Thr Leu Phe Pro Ser Gln His Ser

130

135

140

His Ile Asn Ser Thr

145

<210> 79

<211> 30

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: an artificially synthesized oligo-cap linker sequence

<400> 79

agcaucgagu cggccuuguu ggccuacugg

30

<210> 80

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: an artificially synthesized oligo(dT) primer sequence

<400> 80

gcggctgaag acggcctatg tggccttttt tttttttttt tt

42

<210> 81

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: an artificially synthesized primer sequence

<400> 81

agcatcgagt cggccttggt g

21

<210> 82

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: an artificially synthesized primer sequence

<400> 82

gcggctgaag acggcctatg t

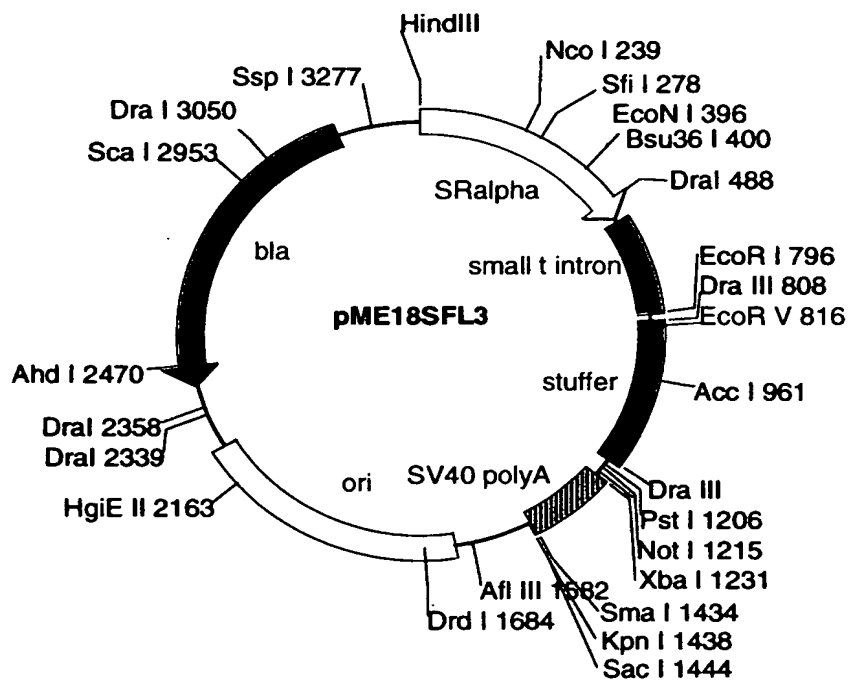
21

【図面の簡単な説明】

【図 1】 pME18SFL3のベクターのマップ

【書類名】 図面

【図 1】



【書類名】 要約書

【要約】

【課題】 新規な全長cDNAの提供。

【解決手段】 ヒトに由来する 3 9 のcDNAを単離した。そしてこのcDNAの全長塩基配列、並びにこの塩基配列によってコードされるアミノ酸配列を明らかにした。本発明のcDNAは全長であるため、翻訳開始点を含み、蛋白質の機能解析において有用な情報を与える。

【選択図】 なし

特願 2 0 0 3 - 1 3 1 4 5 2

出 願 人 履 歴 情 報

識別番号 [5 0 2 2 3 5 5 2 2]

1. 変更年月日	2 0 0 2 年 6 月 2 8 日
[変更理由]	新規登録
住 所	東京都港区西新橋二丁目 3 番 9 号
氏 名	バイオテクノロジー開発技術研究組合